Pertanika J. Sci. & Technol. 32 (5): 2165 - 2191 (2024)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

*Review Article*

# Exploring Probability of Detection (POD) Analysis in Nondestructive Testing: A Comprehensive Review and Potential Applications in Phased Array Ultrasonic Corrosion Mapping

**Jan Lean Tai[1], Mohamed Thariq Hameed Sultan[1,2,3]*, Farah Syazwani Shahar[1], Noorfaizal Yidris[1], Adi Azriff Basri[1] and Ain Umaira Md Shah[1]**

[1]*Department of Aerospace Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[2]*Laboratory of Biocomposite Technology, Institute of Tropical Forest and Forest Product (INTROP), University Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia*
[3]*Aerospace Malaysia Innovation Centre [944751-A], Prime Minister's Department, MIGHT Partnership Hub, Jalan Impact, 63600 Cyberjaya, Selangor Darul Ehsan, Malaysia*

## ABSTRACT

In nondestructive testing (NDT), ensuring defect detection, measurement accuracy, and reliability guarantees various components' structural integrity and safety. The Probability of Detection (POD) concept has emerged as a fundamental measure of the effectiveness of an inspection technique in identifying defects. Since NDT plays a crucial role in aerospace, manufacturing, and infrastructure industries, enhancing POD has become critical. POD refers to the likelihood that a flaw or defect of a certain size will be detected using the NDT technique. The "â versus a" and the "hit/miss" methods are particularly notable among the commonly employed POD estimation methods. The POD curve is determined based on crack size measurements in the "â versus a" approach, typically used in ultrasonic testing. On the other hand, the "hit/miss" method establishes the POD curve by analysing binary outcomes, where a "hit" signifies successful detection and a "miss" denotes detection failure. This review focuses on POD in the context of NDT, specifically in phased array ultrasonic corrosion mapping (PAUCM), to uncover current uncertainty parameters and explore

ARTICLE INFO

*Article history:*
Received: 10 January 2024
Accepted: 12 April 2024
Published: 15 August 2024

DOI: https://doi.org/10.47836/pjst.32.5.14

*E-mail addresses*:
taijanlean2008@hotmail.com (Jan Lean Tai)
thariq@upm.edu.my (Mohamed Thariq Hameed Sultan)
farahsyazwani@upm.edu.my (Farah Syazwani Shahar)
nyidris@upm.edu.my (Noorfaizal Yidris)
adiazriff@upm.edu.my (Adi Azriff Basri)
ainumaira@upm.edu.my (Ain Umaira Md Shah)
* Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

an innovative avenue for enhancing POD assessment by incorporating the material surface temperature as an additional parameter.

*Keywords:* â versus a, hit/miss, model-assisted probability of detection, nondestructive testing, phased array ultrasonic testing, probability of detection

## INTRODUCTION

In nondestructive testing (NDT), ensuring defect detection and measurement accuracy and reliability is crucial for guaranteeing various components' and structures' structural integrity and safety. Probability of detection (POD) provides a quantitative measure of the effectiveness of an inspection technique for defect identification. POD has emerged as a fundamental concept in the aerospace, manufacturing, and infrastructure industries.

The POD measures the likelihood that an NDT technique can detect flaws of a specific size. It is a key performance indicator influencing maintenance, quality control, and risk management decisions. POD analysis empowers engineers, inspectors, and decision-makers to make informed choices based on confidence in defect detection outcomes. The significance of POD lies in bridging the gap between theoretical expectations and real-world applications, ensuring that inspection methodologies align with the safety and quality requirements. Exploring POD is crucial for maintaining inspection accuracy as technology advances in the NDT landscape.

This review aims to comprehensively explore the concept of POD in the context of NDT, specifically emphasising its application in phased-array ultrasonic corrosion mapping (PAUCM). The primary focus was to investigate the intricacies of POD, particularly concerning PAUCM, to uncover the current uncertainty parameters associated with this technique. Furthermore, this review aims to explore an innovative avenue for enhancing POD assessment by incorporating material surface temperature as an additional parameter.

However, owing to the apparent dearth of existing studies in this specific niche, the focus was redirected toward the broader realm of POD assessment within the context of phased-array ultrasonic testing (PAUT). Despite its distinct applications, PAUT is commonly utilised with an angle beam for weld inspection, and the PAUCM employs a straight beam for corrosion detection. The two techniques have the same technology and equipment; there is also a similar mode of operation and the presentation of defects. Therefore, this review broadly explores historical processes, theoretical foundations, methodological strategies, illuminating case studies, and emerging trajectories in the field of POD assessment with a particular focus on the context of PAUT.

This review analyses in-depth contributions from researchers, industry experts, and practitioners to help comprehensively understand the evolution, challenges, and potential opportunities for POD assessment within the PAUT field.

Essentially, this endeavour intends to highlight the critical role of POD as a vital tool in maintaining the reliability and credibility of NDT practices, especially within the PAUT framework. Furthermore, investigating the potential role of the material surface temperature as a POD enhancer will pave the way for innovations that refine the accuracy and effectiveness of PAUCM in corrosion detection.

## PROBABILITY OF DETECTION

POD analysis is a fundamental task in assessing the ability of NDT techniques to detect defects. This comprehensive review aims to outline the intricacies of POD analysis and trace its developmental trajectory, methodology, and significance in protecting material and structural integrity.

Meeker (2000) traced the origins of POD to the 1970s. However, real advancement and widespread adoption began after 2000, as evidenced by a literature review of 90 articles. In the mid-1970s, an estimation of the probability of flaw detection based on specific flaw sizes using uniform POD assumptions was attempted. Initially, binomial distribution methods were used for the estimation. However, this approach proved inadequate, as researchers discovered the multifaceted behaviour of POD curves (Georgiou, 2007).

The NDT community shifted towards more intricate models in the 1980s, which can capture the relationship between POD and flaw size. Log-logistic and 'log-normal' distributions are now utilised as sophisticated approximations of POD behaviour, illustrating the growing comprehension of this intricate phenomenon (Annis et al., 2015a).

The three-parameter model expands upon the two-parameter model by introducing the Signal Amplitude Distribution (SAD), the POD, and the Signal-to-Noise Ratio (SNR). This inclusion offers a more comprehensive representation of the detection process, recognising that flaws may exhibit various signal amplitudes. This change began a new era in POD analysis, leading to more accurate and relevant estimates (Knopp & Zeng, 2013).

Integrating simulation methodologies and advanced statistical techniques offers a more comprehensive and accurate approach to estimating the POD, accounting for a wide range of complex variables and uncertainties. Simulation-based approaches, such as Model-Assisted Probability of Detection (MAPOD), gained prominence during this period. The MAPOD approaches leverage computer simulations to replicate real-world inspection scenarios, assess the probability of flaw detection in different materials with varying sizes and characteristics, and generate realistic data that reflects the complexities of actual NDT operations (Rentala et al., 2018).

MAPOD has gained significant traction since its inception by the MAPOD Working Group in 2004. Its widespread adoption can be attributed to its capability to simulate NDT data for POD assessment, which leads to substantial resource savings (Dominguez et al., 2012).

Simulations have explored multiple variables such as material properties, defect geometries, inspection configurations, and environmental conditions. This holistic approach has enabled researchers to identify optimal parameters and strategies for flaw detection and quantifying uncertainties associated with different scenarios.

In 2009, the US Department of Defence published a notable handbook, MIL-HDBK-1823A, titled "Nondestructive Evaluation System Reliability Assessment." This handbook thoroughly explains POD applications and is frequently cited in various POD research articles, indicating its widespread circulation and influence in the field (DOD, 2009).

Advanced statistical techniques, including Bayesian methods, have been integrated into simulation-based approaches to enhance the accuracy and reliability of POD predictions. By combining the simulation results with statistical analyses, researchers can generate more robust POD curves that account for the variability and uncertainties introduced during inspection (Dominguez et al., 2012). POD estimation methods fall under two primary categories: the "â versus a" and the "hit/miss" methods.

POD computing and evaluation have witnessed the emergence of various techniques over time. In addition to Bayesian methods, the Monte Carlo simulation approach is another commonly employed technique. This approach entails generating random samples from probability distributions and simulating the inspection process to estimate the POD. Monte Carlo simulation accounts for uncertainties, such as defect variability and measurement error, and precisely estimates POD and associated uncertainties (Abdelli et al., 2019).

The 29/29 method calculates the POD using a set of 29 flawed samples. The inspection results are then meticulously analysed to determine the proportion of successfully detected flaws. If all 29 flawed samples were accurately identified, the POD is considered 100% (Bato et al., 2017).

Maximum Likelihood Estimation (MLE) is an established statistical method for determining the parameters of a probability distribution based on observed data. The objective of MLE is to identify the parameter values that provide the highest likelihood of observing the actual detection outcomes, given a hypothesised model. Nevertheless, MLE-based approaches have been known to have certain limitations, such as convergence issues with algorithms. Despite the findings suggesting an increase in POD with crack length, there may be instances in which operators fail to identify large crack sizes, as evidenced by preliminary research (Poudel et al., 2022).

Spies and Rieder (2018) employed the Rayleigh-Rice model to develop MAPOD. The Rayleigh component represents noise, whereas the Rice component represents the signal. This model is particularly relevant in ultrasonic testing because flaw signals are commonly concealed within background noise. By separately characterising the signal and noise components and subsequently combining them, the Rayleigh-Rice model assesses the detectability of flaws.

After thoroughly examining extensive literature, it is apparent that the “â versus a” and binary “hit/miss” methods are the most widely used and central to this article’s focus.

In the “â versus a” approach, such as in ultrasonic testing, the echo amplitude “â” is proportional to the crack size measurement “a”. The objective was to establish a decision threshold “â” that maximised crack detection while minimising false alarms attributed to noise. To address the variability in the “â versus a” relationship, ASTM-E3023 and MIL-HDBK-1823A employ linear functions to assign “â” to “a” and incorporate prediction intervals to factor in noise and statistical uncertainty (Virkkunen et al., 2019).

The accuracy of the output in POD is affected by the amount of the input data. Inadequate data may lead to bias. For instance, MIL-HDBK-1823A recommends analysing at least 40 representative defect data points for “â versus a” and signal strength “â” measurements, as well as crack sizes “a” (Carboni & Cantini, 2016). Conversely, the “hit/miss” method determines the POD curve based on binary outcomes, with “hit” indicating successful detection and “miss” indicating failure. This “hit/miss” technique requires a larger dataset, with a minimum of 60 data sets, to ensure an unbiased and reliable POD curve (Virkkunen et al., 2019).

The receiver operating characteristic (ROC) is a vital metric for binary decisions (“hit/miss” technique). It illustrates four potential prediction outcomes in the binary decision scenarios, particularly for defect presence/absence determinations (Topp & Strothmann, 2021).

- True Positive (TP): When the system correctly spots a problem in a sample with an issue.
- True Negative (TN): When the system correctly says everything is fine with a sample that is indeed problem-free.
- False Positive (FP): When the system mistakenly thinks a problem in a sample is okay. It is like a false alarm.
- False Negative (FN): When the system misses a real problem in a sample, failing to identify it. It is like overlooking an actual issue.

Reliability in NDT is often expressed in the defect size, which has a 90% probability of being detected, referred to as “a90.” This value is presented with a 95% confidence interval to address statistical uncertainty denoted as “a90/95” (Annis, 2014).

Design of Experiment POD (DOEPOD) is a methodology pioneered by the National Aeronautics and Space Administration (NASA). This approach builds on prior NASA work on POD, based on binomial distribution, by introducing the concept of a90/95 (Poudel et al., 2022). Virkkunen et al. (2019) utilised the DOEPOD model, which extends the binomial perspective of hit/miss data. The primary motivation behind the DOEPOD model is to address the limitation of assuming POD as a function of flaw size following a specific model, as in MAPOD estimation.

DOEPOD aims to provide an efficient and accurate methodology for estimating the observed POD and confidence limits for both hit/miss and signal amplitude testing scenarios. Unlike MAPOD, DOEPOD does not assume that the prescribed POD functions, such as logarithmic or similar, are adequate across a wide range of discontinuity sizes and testing system technologies. Therefore, multi-parameter curve fitting or model optimisation approaches are unnecessary to generate a POD curve (Generazio, 2009).

The various methods for calculating POD have their unique efficacy, and the application of POD may vary depending on the scope of application. Table 1 concisely compares the different POD computing methods, highlighting their advantages and limitations.

Table 1
*POD computing and evaluation method comparison*

| POD method | Advantage | Limitation |
|---|---|---|
| Binary Hit/ Miss | • Easy to understand, and quick assessments are appropriate.<br>• Provide clear-cut results for making decisions. | • Neglecting subtleties and uncertainties, the detection procedure is oversimplified.<br>• Does not offer a probabilistic framework for quantifying uncertainty. |
| â versus a | • Provides a methodology based on probability to estimate POD and related uncertainty.<br>• Provides for the estimation of several parameters, increasing adaptability. | • Demands rigorous consideration when choosing the model's parameters and assumptions.<br>• Complicated models or high-dimensional parameter spaces could need a lot of processing power. |
| Bayesian Approaches | • Flexibility in handling complex models and incorporating prior information.<br>• Provides a structure for assessing uncertainty and drawing probabilistic conclusions. | • Computationally demanding, particularly for models with extensive parameter spaces.<br>• The inclusion of subjectivity in the specification of prior distributions may result in the introduction of bias. |
| Monte Carlo Simulation | • POD estimations that are realistic and include related uncertainties.<br>• Ideal for intricate examination situations when there are multiple sources of uncertainty. | • Depends on presumptions of the model's parameters and underlying probability distributions.<br>• Computationally costly, especially when doing large-scale models. |
| 29/29 Method | • Straightforward approach.<br>• Suitable for routine inspections, this system is simple to use and understand. | • Its breadth is restricted because it is based on a predetermined set of faulty samples.<br>• It might not fully represent the variety and ambiguities in actual inspection situations. |
| Maximum Likelihood Estimation (MLE) | • Statistically sound methodology for estimating parameters.<br>• Effective model parameter estimation from observable data. | • Vulnerable to problems with convergence, especially with limited data or sophisticated models.<br>• Assumes that the process of generating the data is accurately represented by the model that is being fitted. |
| Rayleigh-Rice Method | • Particularly ultrasonic testing, which improves the evaluation of defect detectability.<br>• Accuracy is increased by characterising the signal and noise components separately. | • Makes assumptions regarding the distributions of the signal and noise.<br>• Restricted application to different inspection settings or NDT techniques. |

## POD Application in NDT Methods

The concept of POD revolves around quantifying the likelihood of successfully detecting a flaw of size 'a' through a probabilistic function known as POD. This function serves a twofold purpose: it measures the efficacy of NDT methods in identifying such flaws and contributes to developing risk-based maintenance strategies (Yusa et al., 2016).

The integration of POD analysis is essential to ensure the credibility of the NDT inspection procedures. Its fundamental role in validating the dependability of inspection methods is widely recognised, and it is commonly mandated as an integral component of qualification projects, particularly in safety-critical sectors such as aeronautics (Bato et al., 2020). Not all discontinuities can be classified as harmful defects, and it is crucial to assess the size of the discontinuity to determine whether detection is required because small discontinuities within thick structures may not be detrimental.

### *Ultrasonic Testing (UT) and Phased Array Ultrasonic Testing (PAUT)*

Ultrasonic Testing (UT) techniques are essential for NDT because they can detect internal structures and defects in materials. UT has been used for POD analysis and has gained significant attention for improving the reliability of defect detection. POD analysis originated in the 1970s; however, its development accelerated after 2000, with a notable presentation of diverse defects and POD in UT (Meeker, 2000). Exploration of depth as a parameter for model predictions has also begun. POD curves are constructed based on the defect length, depth, orientation, defect type, shape, operator differences, and inspection environment, which can influence inspection accuracy (Subair et al., 2014).

Current methods for determining the probability of detecting the defect length, size, or depth through POD curves involve model simulations, expert insights, experimental trials, or their combinations. Human elements play a crucial role, particularly in inspection methods such as UT and Radiography testing (RT), where human judgment is essential. Automated NDT methods and computer-aided calculations have been employed to address the problem of inconsistencies at actual construction sites (Wall et al., 2009).

Kojima et al. (2019) explored the effect of human factors on POD parameters during ultrasonic inspection. They found that certified and uncertified inspectors had a similar failure risk when detecting stress corrosion cracking in stainless steel pipes.

Flat-bottom holes (FBH) and side-drilled holes (SDH) are commonly used for calibration during ultrasonic testing. FBHs are easy to manufacture and mimic various defects; however, their circular shape may not match them (Stubbs, 2005). An experiment was conducted to detect fatigue cracks and POD using a 0.5 mm diameter and 5 mm high FBH. Actual FBH dimensions represented "a" and ultrasonic defect echo amplitude represented "â" Statistical analysis determined the best linear fit between "a" and "â" (Rentala et al., 2016).

This approach optimises the probability of detecting common in-service fatigue cracks (Carboni & Cantini, 2016). Apart from the manufacturing phase NDT, periodic inspections are crucial for identifying deteriorated structures in service. In-service inspections are more challenging owing to the complexity of defects. Fatigue cracks are a-frequent problem when structures or equipment are in regular use, and POD can aid in fatigue life inspection. The Probability of Failure (POF) can be estimated using ultrasonic-detected defect data and actual defect data (Guan et al., 2014).

PAUT utilises POD curves based on binary 'hit/miss' data but does not account for defect location and dimension accuracy on welds. Consequently, a follow-up "â versus a" analysis was conducted, enhancing accuracy. The ongoing research has been extended to stainless steel and dissimilar materials welds by incorporating real defects and artificial Electrical Discharge Machining (EDM) notches (Kurz et al., 2012; Kurz et al., 2013). One practical application of POD studies lies in PAUT, mainly using reference blocks made from composite materials with FBHs. These studies aim to gauge the reliability and capabilities of PAUT in detecting flaws such as FBHs in composite materials, offering insights into inspection technique performance and sensitivity (Dominguez et al., 2016).

In recent years, researchers have emphasised the integration of the Total Focusing Method (TFM) and Full Matrix Capture (FMC) to enhance the imaging capabilities of PAUT significantly. The TFM technique synthesises multiple ultrasonic waves captured by an array of transducers. Subsequently, the received signals are processed, and the ultrasonic energy is focused on specific points within the material, generating high-resolution images with improved defect detection and characterisation (Caulder, 2018). FMC was developed to address the limitations of conventional PAUT data acquisition methods, which typically capture only a subset of the available ultrasonic data. This approach records the complete set of ultrasonic signals captured by the transducer array, providing a comprehensive dataset that can be utilised for various post-processing techniques, including TFM.

The increasing popularity of TFM and FMC has led to research efforts to refine their underlying algorithms (Zhao et al., 2023). The TFM/FMC algorithm consists of four main steps: (1) data acquisition, which involves obtaining raw ultrasonic data from transducers; (2) signal processing, where noise is removed from the raw data, and system imperfections are corrected; (3) beamforming, in which the processed signals are combined to focus ultrasonic energy on a specific point within the material, and (4) image reconstruction, where the focused signals are utilised to generate a high-resolution image.

Advancements in signal processing, beamforming, and image reconstruction have enhanced performance and reduced computational time for these algorithms. Moreover, machine learning and artificial intelligence techniques are being investigated to further enhance the capabilities of TFM and FMC in PAUT (He et al., 2024). These advanced algorithms enable high-resolution imaging and accurate defect quantification, ultimately improving the overall effectiveness of the inspection process.

POD is a crucial performance metric for TFM and FMC techniques. It is influenced by signal-to-noise ratio, spatial resolution, defect orientation, and defect size (Bajgholi et al., 2023). The ongoing development and enhancement of TFM and FMC algorithms improve POD in PAUT, ultimately leading to more accurate and reliable defect detection and characterisation in various industrial applications.

### *Eddy Current Testing (ECT)*

Eddy Current Testing (ECT) is modelled using complex electromagnetic equations, such as Maxwell's, which are challenging to solve analytically. A numerical approach is preferred for accurate modelling (Abdelli et al., 2019). During the ECT of metallic structures, the probe induces eddy currents in the material, and changes in the coil impedance are detected as the probe traverses a surface or near-surface crack. These changes are measurable parameters and typically increase with the defect size.

A POD demonstration test was conducted to evaluate the performance of the ECT system using standardised specimens with known crack sizes and distributions. The procedure involved scanning the specimens to gather the ECT response data for specific crack sizes, and the POD curve was developed using a two-step analysis:

1. Establishing "â versus a" Relationship: This step employs advanced regression techniques to establishes a mathematical link between the measured EC signal response "â" and actual crack size "a," accounting for factors such as depth and inspection process variability (Zhu et al., 2018).
2. Constructing the POD Curve: Building on normal probability theory, this phase involves creating the POD curve, depicting the probability of detecting a flaw of size "a" based on measured EC signal response "â". This curve gauges the reliability of the EC inspection system for detecting various flaw sizes.

While the "â versus a" relationship provides average ECT signal responses for specific crack sizes, variations can occur in measured responses even for identically sized cracks due to physical attributes of flaws like depth and fluctuations in the ECT process. Statistical techniques can be used to evaluate the capability and reliability of ECT for detecting cracks and defects (Brown, 2009).

Repeated stresses can lead to fatigue-induced cracks, even below a material's breaking point, such as the circular crack growth around a hole. A method called "bolt-hole eddy current" (BHEC) uses a sensor inserted into a hole to detect changes indicative of cracks or issues (Underhill et al., 2018).

Similarly, Underhill and Krause (2011) inspected fatigue cracks in aluminium bolt-holes, generating POD using the crack depth and length as uncertain parameters. A follow-up study by Underhill and Krause (2016) examined corner cracks with 45 EDM notches and 72 fatigue cracks using BHEC.

Finland secures nuclear waste using copper canisters and ensures that their leak-free integrity is crucial. Kanzler et al. (2019) constructed a POD curve for a canister using the ECT method.

However, Xu et al. (2023) emphasised that POD would decrease due to the decrease in eddy current density, leading to the deterioration of SNR.

### *Radiography Testing (RT)*

Weld defects exhibit diverse shapes, which makes it challenging to establish reliable and efficient POD assessments (Kanzler & Müller, 2016a). RT features a dynamic threshold adjusted based on defect size. This method tailors the detection thresholds according to the defect dimensions. However, acquiring sufficient data for POD assessment can be resource-intensive. Weld defects exhibit diverse shapes, which makes it challenging to establish reliable and efficient POD assessments (Kanzler & Müller, 2016a).

Innovative methodologies are required to address the complexity of utilising the actual defects for POD calculations. One such strategy is an indication size-dependent evaluation that addresses two key defect parameters. However, this approach requires a large amount of data. A smoothing algorithm is introduced to enhance the evaluation accuracy, which considers physical characteristics and defect detection capabilities (Kanzler & Müller, 2016b).

The Bayesian approach is useful for calculating POD curves, especially in situations with few defects. This method is essential when detailed results are required, such as nuclear fuel disposal canisters. The Bayesian approach was used to derive POD curves in this context because it meets the strict safety requirements for nuclear applications, and few actual defects are owing to high-quality production techniques. This method combines available data, expert knowledge, and statistical reasoning to provide reliable insights into the performance and capabilities of RT (Kanzler et al., 2012).

The influence of RT techniques on POD evaluation for planar flaws, such as cracks, is affected by various factors, including the flaw orientation relative to the beam direction, human factors, application conditions, accessibility, equipment sensitivity and resolution, manufacturing processes, and material properties. Equations were used to construct POD curves for cracks using varying parameters, such as grain diameter, thickness sensitivity percentage, crack width, and number of orientations. The results showed that increasing the number of orientations and crack width increased POD, whereas higher sensitivity values and larger object diameters decreased POD (Ghose, 2013).

Computed Radiography (CR) uses reusable storage phosphor plates instead of films for industrial Radiography. Research has shown that medium-resolution CR systems perform better for flaw detection at certain dose thresholds, which affect the computed POD. It suggests that medium-resolution CR systems are more effective for detecting flaws detection at these dose levels (Mohr & Willems, 2008).

It is crucial to compare the POD of various defect detection algorithms using artificial X-ray Computed Tomography (XCT) data from industrial specimens. Artificial XCT data were generated through numerical XCT simulations, which enabled the controlled incorporation of specific defects or pores at predefined locations (Yosifov et al., 2023). XCT, a vital volumetric imaging technique, is widely used in X-ray-based digital radiography (XDR) and POD computation. This approach often involves creating specimens with distinct artificial defects to capture the shape and size variations.

XCT simulations have applications in the biomedical and material science sectors for virtual radiographic testing optimisation and forecasting NDT systems' reliability using ray-tracing algorithms similar to SimCT to generate radiographic images (Yosifov et al., 2022).

A study by Kim et al. (2021) carried out an XCT experiment where NDT signals' â' were compared with direct measurements of the actual property value. Statistical modelling accounted for the inherent noise in the NDT signal. Notably, the same types of flaws may not consistently produce the same signal, and even flaws of identical size might result in different signals. The estimation of a POD curve is grounded in NDT measurements taken from flaws of different sizes. Initially, this method was employed to evaluate flaws during the operational lifespan of a component.

### *Structural Health Monitoring (SHM) and Guided Wave (GW)*

Structural Health Monitoring (SHM) involves embedding sensors in a system to continuously evaluate structural health throughout its operational life. Guided waves, or Lamb waves, are sequentially emitted and received by the sensors to generate a comprehensive structure scan. Comparing these scans with a baseline scan captured when the structure was pristine provides a visual representation of the structural condition (Calmon et al., 2019).

These waves effectively identify surface and internal structural anomalies such as delamination, holes, cracks, corrosion, and wear in lap joints. Assessment of NDT capabilities often uses the Berens model, which employs POD curves to illustrate the likelihood of flaw detection based on size (Gianneo et al., 2016a). However, POD curves may exhibit nonlinear patterns with respect to the crack size, reflecting fluctuations supported by numerical simulations and empirical data (Gianneo et al., 2016b).

Forsyth (2016) proposed a novel approach using a single specimen with a growing crack to address the variability in crack responses and the impact of repeated inspections on the POD estimation. Forsyth (2016) emphasised that in many POD studies, the primary source of variability is not the measurement process but rather the diverse responses among cracks of similar sizes. It implies that reducing the sampling may lead to inaccurate POD estimates. Although treating repeated inspections as independent events to enhance POD is appealing, it contradicts the literature.

NDT and SHM share similarities, but their sensor setups affect POD curve interpretation. Portable sensor arrays are fundamental to NDT methods, such as ultrasonic wave testing, whereas guided wave monitoring in SHM employs permanently fixed transducers. POD curves help to evaluate the damage-detection capabilities within a predefined setup and damage location for SHM systems (Bayoumi et al., 2021).

Tschoke et al. (2021) project focussed on creating an SHM system for safety-critical components made from carbon fibre-reinforced polymer (CFRP) in the automotive sector. POD varies based on the defect size, as shown in the POD curves. However, the regulatory framework for SHM is still evolving, and a significant challenge is the lack of reliable methods for assessing POD in SHM systems.

### *POD Application in Diverse NDT Methods*

The UT and ECT are prominent NDT techniques. However, NDT encompasses diverse methods, each significant in defect detection and characterisation.

Fluorescent Penetrant Testing (PT) is a versatile technique used throughout the various stages of manufacturing to identify surface cracks. The objective was to detect linear indications using POD curves that encapsulate the inspector's ability to detect specific crack sizes. The analysis categorises outcomes as hits, misses, or false calls and quantifies an inspector's proficiency (Herberich, 2009).

A comprehensive case study on PT was delved into a POD investigation across 27 titanium samples involving multiple inspectors. The resulting POD curve provides insights into crack detection effectiveness, accounting for factors such as inspector performance and testing conditions (Caturano et al., 2009).

Acoustic Emission (AE) monitoring is a valuable technique for detecting wire failures because it captures audible sounds from material defects. AE technology provides a comprehensive solution for identifying cable breaks, fatigue cracks, and corrosion (Lembersky et al., 2012). However, challenges arise due to the unique nature of diagnostic signals and source dynamics, making distinguishing genuine AE signals from noise difficult. A holistic approach, including stress stimuli, source behaviour, wave propagation, sensor sensitivity, and detection threshold, provides insights into the POD (Hossain et al., 2013).

Incorporating Infrared thermography (IR) into CFRP materials involves creating POD curves through experimental methods. Testing CFRP samples with known defects provided a statistical analysis, offering insights into the technique's defect identification performance (Peeters et al., 2018).

Pulsed Thermography is effective for materials exhibiting rapid heat diffusion. POD analysis aids in quantifying defects and assessing accuracy by stimulating materials with energy bursts and capturing thermal data (Accardi et al., 2023).

## Model-Assisted Probability of Detection

Recent advancements in POD analysis have involved incorporating computational models and simulations to improve defect detection. This study examines the evolution of MAPOD by tracing its development from traditional POD analysis to advanced simulation techniques. MAPOD originated in the early 2000s when scholars first recognised the potential of computer-aided models in ultrasonic corrosion mapping for POD estimations (Burch et al., 2005).

The MAPOD working group was established in the US in 2004, and later, similar initiatives emerged in Europe, such as SISTAE in France and PICASSO (Dominguez et al., 2016). These efforts aimed to bridge the gap between mathematical models and experimental data, thereby allowing the prediction of a broader range of defect types and inspection scenarios.

As POD models evolve, diverse mathematical frameworks have emerged to support the MAPOD analyses. MAPOD simulates data using statistical or finite element methodologies, reducing reliance on resource-intensive experimental data (Wright, 2016). Conventional POD analyses are limited by their reliance on a single parameter; however, Yusa et al. (2018) advocated for a multi-parameter POD model facilitated by numerical simulations. Baskaran et al. (2021) extended the MAPOD paradigm to ECT, harnessing multiple flaw response signals, including the defect length and coil impedance at different frequencies.

Although MAPOD offers efficiency, validating the resulting probabilistic POD curves is also essential. A study by Le Gratiet et al. (2017) compared POD curves generated using four methods: Behrens, binomial-Barens, polynomial chaos, and kriging. The study found subtle differences in the lower 95% bounds of the a90/95 estimates, highlighting the complexity of the detectable defect size estimation.

Rodat et al. (2017) applied the MAPOD methodology to the ultrasonic inspection of composite materials, incorporating input vectors such as material thickness, FBH diameter, and surface defect depth (FBH depth). The resultant output vector captured defect characteristics.

### *Mh1823 POD Software*

The MIL-HBDK-1823A manual provides a comprehensive guide for constructing POD studies (DOD, 2009). It also includes a helpful Mh1823 POD software download guide built on the R statistical and graphics engine, which can be accessed on the Statistical Engineering website (https://www.r-project.org/).

Tschoke et al. (2021) demonstrated the automotive industry's application of POD analyses in creating SHM systems for CFRP components. They emphasised the international recognition of the MIL-HDBK manual and Mh1823 software as well-established standards for conducting these evaluations.

Forsyth and Aldrin (2009) provided a practical demonstration for conducting a POD curve study on BHEC using the Mh1823 software. Choi et al. (2022) replaced RT with UT and PAUT and detected volumetric defects through round-robin tests on various materials. POD analysis was performed using the Mh1823 POD analysis software, while a simulation using finite element techniques explored the potential of eddy current testing for detecting stress corrosion cracking signals (Yusa, 2017). The resulting POD curves were generated using the R software.

Kurz et al. (2012) documented using Mh1823 in the PAUT domain. TFM, a synthetic focusing technique, was integrated with Mh1823. It uses full-matrix capture to capture fundamental ultrasonic signals, enhances defect boundary delineation, generates coherent signals, and mitigates noncorrelation artefacts (Bajgholi et al., 2023).

Ground-penetrating radar (GPR) is used in civil engineering to detect rebar and tendon ducts within concrete structures. Its effectiveness is highlighted by its ability to uncover the subsurface features. Mh1823 software was employed for uncertainty analysis to evaluate GPR's reliability of the GPR. This software is a reliable tool for determining the precision of GPR in identifying subsurface features within concrete structures (Feistkorn & Taffe, 2014). Remarkably, Mh1823 has additional applications beyond NDT, including predicting driving behaviour (Ameyaw et al., 2019).

### *Computational Intelligence for Visual Applications (CIVA) Software*

CIVA began in the early 1990s and evolved into a comprehensive toolkit with modules for various NDT methods, including UT, GW, ECT, RT, and CT. It is now an all-encompassing solution for simulating and analysing diverse NDT techniques across industries (Foucher et al., 2018). CIVA enables the generation of POD curves more efficiently and cost-effectively based on the precision, consistency, and repeatability of theoretical POD simulations. In pulsed-echo ultrasonic inspection, parameters such as defect length or size are crucial for generating POD curves. CIVA facilitates MAPOD through Monte Carlo simulations and establishes relationships between the input parameters and POD outcomes (Schneider et al., 2012).

CIVA is useful for calculating various uncertainty parameters. Dominguez et al. (2012) show how CIVA can be used to calculate the uncertainty parameters for PAUT. Automation reduces human errors during the inspection process. Parameters such as the water parts, defect angular position, and radial position were used for the uncertainty analysis for the PAUT. The study confirmed that a well-designed PAUT procedure can achieve 90% probability and 95% confidence in detecting a 0.5 mm diameter void.

Ribay et al. (2017) examined the use of CIVA for generating 'hit/miss' studies in the PAUT of centrifugally cast stainless steel pipes. This study focuses on high-attenuation surfaces with coarse grain structures and material thicknesses as crucial parameters for the analysis. Haapalainen and Leskelä (2012) used CIVA to generate defect-size-dependent 'a' vs. signal response 'â' curves to detect service-induced cracks.

Marcotte and Liyanage (2017) combined multiple NDT techniques for inspection. They used ECT and PAUT and obtained promising results for detecting target defects. The study validated the CIVA software and suggested using multi-technique inspection systems.

Dominguez et al. (2010) and Jenson et al. (2011) conducted High-Frequency Eddy Current Testing (HFET) to detect fatigue cracks in titanium alloys. Their study demonstrated a close alignment between simulated POD and experimental POD, bolstering confidence in the CIVA's predictive capabilities.

Similarly, CIVA is instrumental in simulating ECT scenarios. Goursolle et al. (2016) simulated ECT on fatigue cracks in Inconel 718 material, emphasising the influence of parameters such as inspectors, air gap, and frequency settings. Bato et al. (2017) and Bato et al. (2020) highlight how CIVA offers advantages in evaluating the impact of human factors on POD through simulation rather than experimentation. This approach contributes to the enhanced credibility of the model.

After that, CIVA extends its reach to RT. Tisseur et al. (2019) elaborated on CIVA's version 11, which introduced a POD simulation tool within the RT module. This tool caters to scenarios where the radiation source is aligned with the defect or circumferentially misaligned, demonstrating the platform's versatility.

### *Other Software Recommended by Researchers*

Lei et al. (2022) highlighted the significance of various factors in POD analysis, such as test methods, materials, defects, equipment, and human factors. These multifaceted influences are owed to the complexity of POD estimation. They introduced SimSUNDT software, designed by the Chalmers University of Technology in Sweden, to enhance PAUT's capability to replicate intricate inspection scenarios.

Subair et al. (2014) used simulation-driven POD estimation with ABAQUS software to explore pulsed electron ultrasound propagation for stainless steel surface notch detection. They evaluated the impact of various factors, including probe position, incident wave angle, and ultrasound frequency. The authors compared logarithmic scatter plots of simulated and experimental defect response signals using a simulation-experimental correlation approach and validated their findings using linear regression analysis in MATLAB.

Volker et al. (2004) introduced the "POD-generator" software, designed to enhance the structural integrity of pipes and pipelines through corrosion inspections using an ultrasonic technique. The software integrates data from the inspection and degradation models to determine the integrity of the three components. It produces a curve based on POD in specific scenarios, thereby improving the reliability in assessing structural conditions.

Fusing simulated and experimental data enhances the NDT performance evaluation and reduces costs. Gollwitzer et al. (2011) introduced the concept of merging these datasets to enhance the NDT performance estimation. The aRTist tool is the foundation for this approach. SimuPOD provides a user-friendly interface for defining the calculation series and automating the analysis, focusing on POD studies. This innovation streamlines the reliability and effectiveness of the NDT methods and increases their practical applicability.

## Literature Review Summary

NDT techniques must be accurate, reliable, and critical across aerospace, manufacturing, petrochemical, and civil engineering industries. POD analysis is a powerful tool that quantifies inspection reliability by considering the defect size, materials, inspection configuration, and human factors. Linking NDT methods and risk assessment enhances decision-making and improves safety and efficiency.

An exploration encompassing around 70 articles titled “POD” or “probability of detection” from 2000 to 2023 reveals that the concept of POD has garnered extensive application across various NDT methods. The trajectory of its usage has revealed an intriguing pattern shaped by both temporal evolution and technological advancements.

POD originated in the aerospace industry owing to its strict demands. ECT is a crucial tool because of its widespread use in the aerospace industry. UT was among the first to adopt POD. In this review of 70 articles, ECT methods were mentioned in 18 papers, whereas UT was mentioned in 17 papers, totalling half of the corpus.

PAUT closely followed ten papers, indicating its established position in the NDT landscape. However, as the software landscape matured, there was a shift towards more sophisticated analytical capabilities. This is evident from the substantial growth in RT and GW SHM, which will account for 20% of the literature combined. This evolution shows the profound impact of software advancements in facilitating intricate analyses, expanding the POD’s purview to include crucial domains such as SHM.

Despite having a smaller proportion, other NDT methods play a significant role in defect detection. PT and AE accounted for 4%, indicating their values in the field. The IR and other combined methods comprised the remaining portions. This distribution highlights the diverse avenues through which POD affects NDT, as shown in Figure 1.

The articles’ analyses showed the evolution of POD in NDT methods and its correlation with software development. ECT and UT initially had a stronghold, but broader adoption of POD occurred in RT and SHM. POD enhances defect detection and is critical in various industries. The trend of employing POD software is set to expand with the increase in the MAPOD methodology, leading to greater efficiency and cost-effectiveness.

MAPOD is undeniable owing to technological evolution, offering time and cost savings and rapid generation of accurate POD curves using simulation data. It expedites defect detection and assessment while maintaining reliability.

Throughout the reviewed literature, other researchers' utilisation of POD software has been documented in Table 2.

The specific testing method used can influence the parameter variations. Although general factors remain relevant, additional considerations arise in RT, such as the direction

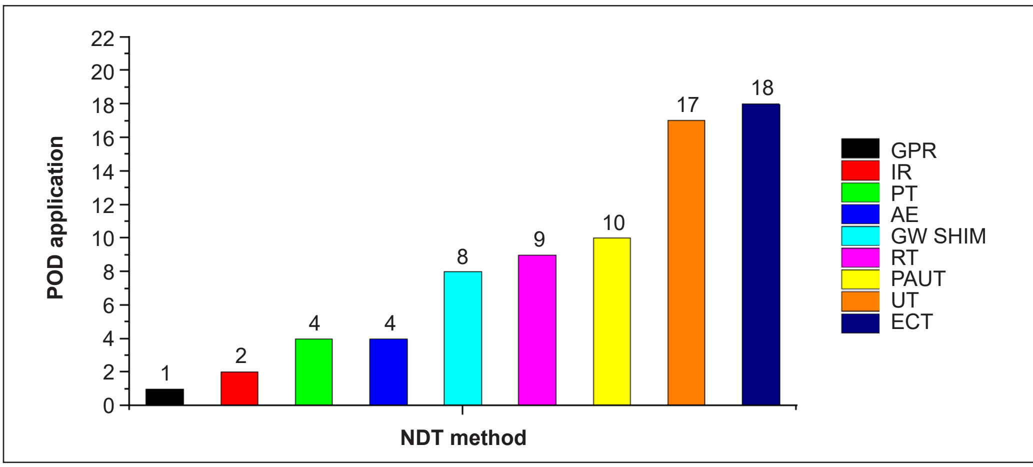


*Figure 1*. Distribution of POD application in different NDT methods

Table 2
*Selected software for POD studies by other researchers*

| Software | Description | Reference |
|---|---|---|
| Mh1823 POD | Is a free software that can be accessed from the Statistical Engineering website (https://statistical-engineering.com/) | Forsyth & Aldrin, 2009; Kurz et al., 2012; Yusa, 2017; Tschoke et al., 2021; Choi et al., 2022; Bajgholi et al., 2023 |
| CIVA | The CIVA software is a versatile commercial tool that extends its utility beyond POD analysis. It encompasses a range of simulation software for various NDT methods. | Dominguez et al., 2010; Jenson et al., 2011; Schneider et al., 2012; Haapalainen & Leskelä, 2012; Dominguez et al., 2012; Dominguez et al., 2016; Goursolle et al., 2016); Ribay et al., 2017; Marcotte & Liyanage, 2017; Bato et al., 2017; Foucher et al., 2018; Calmon et al., 2019; Tisseur et al., 2019; Bato et al., 2020 |
| MATLAB | MATLAB provides various built-in functions and toolboxes for various applications, including mathematics, engineering, physics, finance, image processing, machine learning, and more. | Subair et al., 2014 |
| simSUNDT | It is a simulation software for UT | Lei et al., 2022 |
| aRTist SimuPOD | The aRTist is a computer simulation of both film and digital radiography. SimuPOD is one of the modules. | Gollwitzer et al., 2011 |

of the testing beam and the testing process itself. In the PAUT, the defects' water paths, as well as the angular and radial positions, need to be considered. These method-specific variations highlight the complexity of parameter selection in the POD analysis.

## DISCUSSION

The primary objective of this literature review is to identify existing examples of utilising POD in the context of PAUCM. While PAUCM boasts numerous advantages in corrosion detection, recent studies have highlighted its adaptability for in-line inspections, particularly in elevated surface temperature conditions (Tai et al., 2023). Despite the robust presentation of detection data, integrating POD as a crucial tool for upholding the reliability and credibility of the PAUCM process would provide additional substantiation for the dependability of this application.

Although specific instances of applying POD in the context of PAUCM have not been uncovered, the review has elucidated the widespread utilisation of POD in other NDT methods. Concurrently, it has underscored the significance of uncertain critical parameters as pivotal inputs. By commencing with considerations of defect length and depth, the investigation expanded to encompass various factors, including defect type, size, dimensions, orientation, shape, and location, while also addressing the influence of human factors.

The application of PAUCM could explore an innovative avenue for enhancing POD assessment by incorporating the material surface temperature as an additional parameter, potentially yielding more robust results.

Additionally, this review indicates the existence of two primary POD models: the "hit/miss" model for image-type defects and the "â versus a" model for defects represented in signal amplitude forms, as illustrated in Figures 2 and 3.

PAUCM, a manifestation of PAUT, presents information through A-, B-, C-, and S-scan images, thus providing a comprehensive three-dimensional perspective of the defects.

The A-scan mode resembles the traditional UT mode and displays the ultrasound echo amplitudes. The B, C, and S scans offer essential imaging tools for accurate defect localisation, as shown in Figure 4. The PAUCM ultrasound beam is aligned perpendicular to the test object, like the 0-degree normal probe in the UT.

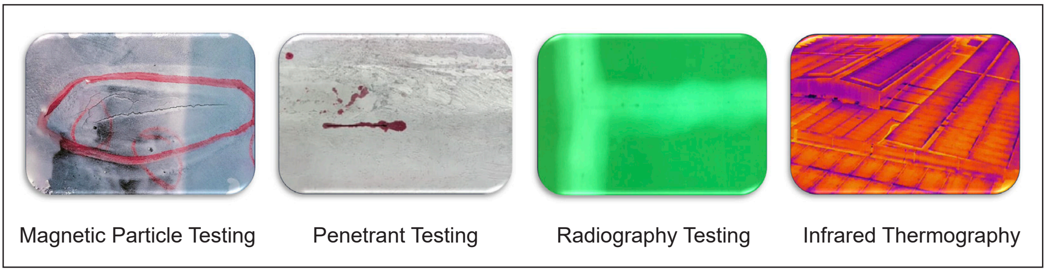


*Figure 2*. Image type NDT methods suitable for "Hit/Miss" POD

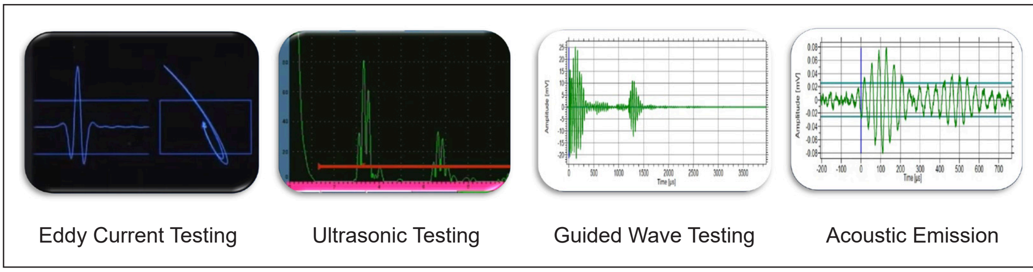


*Figure 3*. Amplitude type NDT methods suitable for "â versus a" POD

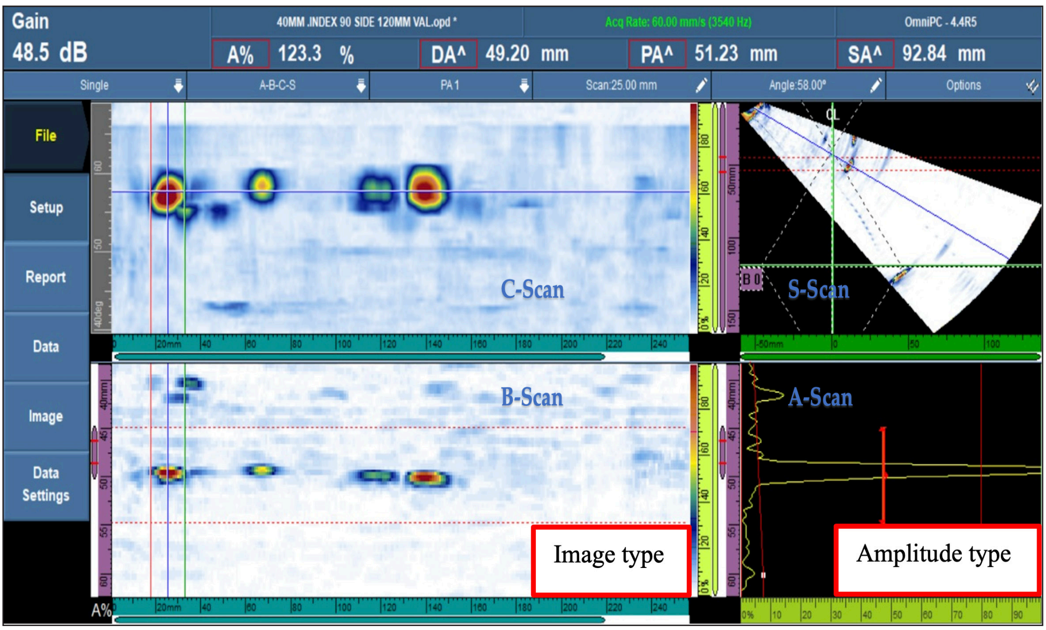


*Figure 4*. PAUT A-, B-, C- and S-scan presentation

PAUCM simultaneously presents information in image type and signal amplitude forms, making it suitable for both primary POD models. This dual representation not only enhances the feasibility of POD but also allows for relative comparisons.

The primary challenge currently lies in the practical application of POD despite a clear understanding of its foundational concepts. While other researchers have employed various POD computing methods, there remains a gap in translating this knowledge into practical implementation. Moving forward, the focus will be on mastering the utilisation of the MH1823 POD software, using the comprehensive guidance outlined in its accompanying handbook. In addition, dedicated experimental studies will be conducted for PAUCM, allowing for the precise collection of data. Subsequent analyses will rigorously examine various uncertainty parameters to construct and compare POD curves carefully, thereby improving the reliability of the assessments.

## CONCLUSION

This comprehensive review highlights the dynamic evolution of POD and its multifaceted applications to various NDT methods. The exploration encompasses the nuanced utilisation of both "hit/miss" and "â versus a" approaches, the emergence and significance of MAPOD, the expedient role of cutting-edge POD software in curve generation, and the exciting prospect of PAUCM within the POD framework.

The journey through this scholarly terrain underscores the remarkable versatility of POD, traversing a broad spectrum of NDT methods and catering to diverse application domains. POD is a unifying metric for assessing the efficacy of defect detection techniques, whether employing UT, ECT, RT, or GW SHM, which mirrors the intricacies of real-world inspection scenarios, where the choice of the NDT method depends on the specific inspection goal and context.

A pivotal insight gleaned from the literature is the significance of parameter selection in the POD analysis. The dichotomy of "red and green apples," as eloquently shared by one of the POD luminaries, encapsulates the essence of this challenge (Annis et al., 2015b). Selecting appropriate uncertainty parameters is critical to ensure that the chosen parameters accurately represent the characteristics of actual defects and align with the intended applications.

Concluding the literature review, it is evident that further exploration of PAUCM is worthwhile. The practical implementation of POD in real-world inspection scenarios remains a vital milestone. This journey necessitates a harmonious interplay between theoretical and practical insights, allowing for the effective integration of POD concepts into the operational landscape of PAUCM.

## ACKNOWLEDGEMENTS

The authors thank Universiti Putra Malaysia for the financial support through Geran Inisiatif Putra Siswazah (GP-IPS) with grant number 9739200. The authors would also like to express their gratitude to the Department of Aerospace Engineering, Faculty of Engineering, Universiti Putra Malaysia, and the Laboratory of Biocomposite Technology, Institute of Tropical Forestry and Forest Product (INTROP), Universiti Putra Malaysia (HICOE) for their close collaboration in this study.

## REFERENCES

Abdelli, D. E., Nguyen, T. T., Clenet, S., & Cheriet, A. (2019). Stochastic metamodel for probability of detection estimation of eddy-current testing problem in random geometric. *IEEE Transactions on Magnetics*, *55*(6), 1–4. https://doi.org/10.1109/TMAG.2019.2893421

Accardi, E. D., Palumbo, D., Errico, V., Fusco, A., Angelastro, A., & Galietti, U. (2023). Analysing the probability of detection of shallow spherical defects by means of pulsed thermography. *Journal of Nondestructive Evaluation*, *42*(1), 1–16. https://doi.org/10.1007/s10921-023-00936-y

Ameyaw, D. A., Deng, Q., & Söffker, D. (2019). Probability of detection (POD)-based metric for evaluation of classifiers used in driving behavior prediction. *Proceedings of the Annual Conference of the Prognostics and Health Management Society, PHM, 11*(1), 1–7. https://doi.org/10.36001/phmconf.2019.v11i1.774

Annis, C. (2014). Influence of sample characteristics on probability of detection curves. *40th Annual Review of Progress in Quantitative Nondestructive Evaluation AIP Conference Proceedings*, *1581*, 2039–2046. https://doi.org/10.1063/1.4865074

Annis, C., Aldrin, J. C., & Sabbagh, H. A. (2015a). Profile likelihood: What to do when maximum probability of detection never gets to one. *Materials Evaluation*, *73*(1), 96–100.

Annis, C., Aldrin, J. C., & Sabbagh, H. A. (2015b). What is missing in nondestructive testing capability evaluation? *Materials and Design*, *73*(1), 44–54.

Bajgholi, M. E., Rousseau, G., Ginzel, E., Thibault, D., & Viens, M. (2023). Total focusing method applied to probability of detection. *International Journal of Advanced Manufacturing Technology*, *126*(7–8), 3637–3647. https://doi.org/10.1007/s00170-023-11328-x

Baskaran, P., Pasadas, D. J., Ramos, H. G., & Ribeiro, A. L. (2021). Integration of multiple response signals into the probability of detection modelling in eddy current NDE of flaws. *NDT and E International*, *118*, Article 102401. https://doi.org/10.1016/j.ndteint.2020.102401

Bato, M. R., Hor, A., Rautureau, A., & Bes, C. (2017). Implementation of a robust methodology to obtain the probability of detection (POD) curves in NDT: Integration of human and ergonomic factors. *LES JOURNÉES COFREND 2017*, 1–16.

Bato, M. R., Hor, A., Rautureau, A., & Bes, C. (2020). Experimental and numerical methodology to obtain the probability of detection in eddy current NDT method. *NDT and E International*, *114*, 1–35. https://doi.org/10.1016/j.ndteint.2020.102300

Bayoumi, A., Minten, T., & Mueller, I. (2021). Determination of detection probability and localization accuracy for a guided wave-based structural health monitoring system on a composite structure. *Applied Mechanics*, *2*(4), 996–1008. https://doi.org/10.3390/applmech2040058

Brown, J. H. (2009). *Probability of detection analysis for eddy current inspection systems*. The American Society for Nondestructive Testing.

Burch, S. F., Stow, B. A., & Wall, M. (2005). Computer modelling for the prediction of the probability of detection of ultrasonic corrosion mapping. *Insight: Non-Destructive Testing and Condition Monitoring*, *47*(12), 761–764. https://doi.org/10.1784/insi.2005.47.12.761

Calmon, P., Mesnil, O., Miorelli, R., Artusi, X., Chapuis, B., & D'Almeida, O. (2019). Model assisted probability of detection for guided wave imaging structural health monitoring. *Proceedings of the 12th International Workshop on Structural Health Monitoring*, *1*, 811–816. https://doi.org/10.12783/shm2019/32190

Carboni, M., & Cantini, S. (2016). Advanced ultrasonic "Probability of detection" curves for designing in-service inspection intervals. *International Journal of Fatigue*, *86*, 77–87. https://doi.org/10.1016/j.ijfatigue.2015.07.018

Caturano, G., Cavaccini, G., Ciliberto, A., Pianese, V., & Fazio, R. (2009). Probability of detection for penetrant testing in industrial environment. In *Applied and industrial mathematics in Italy III* (pp. 186-195). World Scientific. https://doi.org/10.1142/9789814280303_0017

Caulder, A. (2018). Full matrix capture and total focusing method: The next evolution in ultrasonic testing. *Materials Evaluation*, *76*(5), 591–597.

Choi, Y. M., Kang, D., Kim, Y. L., Cho, S., Park, T., & Park, I. K. (2022). Reliability assessment of PAUT technique in Lieu of RT for tube welds in thermal power plant facilities. *Applied Sciences*, *12*(12), Article 5867. https://doi.org/10.3390/app12125867

DOD. (2009). *MIL-HDBK-1823A, Nondestructive evaluation system reliability assessment*. Department of Defense Handbook.

Dominguez, N., Feuillard, V., Jenson, F., & Willaume, P. (2012). Simulation assisted pod of a phased array ultrasonic inspection in manufacturing. *AIP Conference Proceedings*, *1430*(31), 1765–1772. https://doi.org/10.1063/1.4716425

Dominguez, N., Jenson, F., & Defense, E. A. (2010, June 7-11). Simulation assisted POD of a high frequency Eddy current inspection procedure. In *Proceedings of the 10th European Conference on Non-Destructive Testing. European Conference on Non-Destructive Testing* (pp. 1-10). Moscow, Russia.

Dominguez, N., Rodat, D., Guibert, F., Rautureau, A., & Calmon, P. (2016). POD evaluation using simulation: Progress, practice and perspectives regarding human factor. *AIP Conference Proceedings*, *1706*, 3–9. https://doi.org/10.1063/1.4940651

Feistkorn, S., & Taffe, A. (2014). Methods to assess the quality of non-destructive testing in civil engineering using POD and GUM for static calculations of existing structures. *Materialpruefung/Materials Testing*, *56*(7–8), 611–616. https://doi.org/10.3139/120.110602

Forsyth, D. S. (2016). Structural health monitoring and probability of detection estimation. In *AIP Conference Proceedings* (Vol. 1706, No. 1). AIP Publishing. https://doi.org/10.1063/1.4940648

Forsyth, D. S., & Aldrin, J. C. (2009, June 24-26). Build your own POD. In *Proceedings of the 4th European-American Workshop on Reliability of NDE* (pp. 1–8). Berlin, Germany.

Foucher, F., Fernandez, R., Leberre, S., & Calmon, P. (2018). New tools in CIVA for model assisted probability of detection (MAPOD) to support NDE reliability studies. *NDE of Aerospace Materials & Structures 2018*, 32–43.

Generazio, E. R. (2009). Design of experiments for validating probability of detection capability of NDT systems and for qualification of inspectors. *Materials Evaluation*, *67*(6), 730–738.

Georgiou, G. A. (2007). PoD curves, their derivation, applications and limitations. *Insight: Non-Destructive Testing and Condition Monitoring*, *49*(7), 409–414. https://doi.org/10.1784/insi.2007.49.7.409

Ghose, B. (2013). Evaluation of probability of detection (POD) and minimum number of exposures required for detection of planar flaw in cylindrical object by radiographic NDE method. *Asia Pacific Conference on Non-Destructive Testing*, *19*, 1-6.

Gianneo, A., Carboni, M., & Giglio, M. (2016a). Feasibility study of a multi-parameter probability of detection formulation for a lamb waves–based structural health monitoring approach to light alloy aeronautical plates. *Structural Health Monitoring*, 16(2), 225–249. https://doi.org/10.1177/1475921716670841

Gianneo, A., Carboni, M., & Giglio, M. (2016b). Reliability aspects and multi-parameter POD formulation for guided wave based SHM techniques. *19th World Conference on Non-Destructive Testing 2016*, 1–11.

Gollwitzer, C., Bellon, C., Deresch, A., & Ewert, U. (2011). On POD estimations with radiographic simulator aRTist. In *International Symposium on Digital Industrial Radiology and Computed Tomography (No. DGZfP-BB 128 [Tu. 2.3])* (pp. 1-8). Deutsche Gesellschaft für zerstörungsfreie Prüfung eV (DGZfP).

Goursolle, T., Fauret, T., & Juliac, E. (2016, June 13-17). Effect of data amount on probability of detection estimation: Application to Eddy current testing. In *19th World Conference on Non-Destructive Testing 2016* (pp. 1-8). Munich, Germany.

Guan, X., Zhang, J., Zhou, S., Rasselkorde, E. M., & Abbasi, W. (2014). Probabilistic modeling and sizing of embedded flaws in ultrasonic non-destructive inspections for fatigue damage prognostics and structural integrity assessment. *NDT and E International*, *61*, 1–9. https://doi.org/10.1016/j.ndteint.2013.09.003

Haapalainen, J., & Leskelä, E. (2012, April 16-20). Probability of detection simulations for ultrasonic pulse-echo testing. In *18th World Conference on Nondestructive Testing* (pp. 1-5). Durban, South Africa.

He, X., Jiang, X., Guo, J., Xu, L., & Mo, R. (2024). Ultrasonic evaluation of wire-to-terminal joints: integrating XGBoost machine learning with finite element feature analysis. *Nondestructive Testing and Evaluation*, 1–18. https://doi.org/10.1080/10589759.2024.2304265

Herberich, J. (2009). Applying MIL-HDBK-1823 for POD demonstration on a fluorescent penetrant system. *Materials Evaluation*, *67*(3), 293–301.

Hossain, M., Ziehl, P., Yu, J., Caicedo, J., & Matta, F. (2013). *Assessing probability of detection based on acoustic emission associated with fatigue crack extension in steel bridge elements*. The American Society for Nondestructive Testing.

Jenson, F., Iakovleva, E., & Dominguez, N. (2011). Simulation supported POD: Methodology and HFET validation case. *AIP Conference Proceedings*, *1335*, 1573–1580. https://doi.org/10.1063/1.3592117

Kanzler, D., & Müller, C. (2016a, June 13-17). Evaluating RT systems with a new POD approach. In *Proceedings of the 19th World Conference on Non-Destructive Testing* (pp. id-19535). Munich, Germany.

Kanzler, D., & Müller, C. (2016b). How much information do we need? A reflection of the correct use of real defects in POD-evaluations. In *AIP Conference Proceedings* (Vol. 1706, No. 1). AIP Publishing. https://doi.org/10.1063/1.4940652

Kanzler, D., Müller, C., Pitkänen, J., & Ewert, U. (2012, April 16-20). Bayesian approach for the evaluation of the reliability of non-destructive testing methods: Combination of data from artificial and real defects. In *18th world conference on nondestructive testing* (pp. 1-6). Durban, South Africa.

Kanzler, D., Milsch, S., Pavlovic, M., Müller, C., & Pitkänen, J. (2019). Concept of total reliability of NDT methods for inspection of the EB weld of the copper canister used for a long-term storage of spent nuclear fuel. *Structural Integrity and NDE Reliability III Concept*, 1-6.

Kim, F. H., Pintar, A., Obaton, A. F., Fox, J., Tarr, J., & Donmez, A. (2021). Merging experiments and computer simulations in X-ray computed tomography probability of detection analysis of additive manufacturing flaws. *NDT and E International, 119*, Article 102416. https://doi.org/10.1016/j.ndteint.2021.102416

Knopp, J. S., & Zeng, L. (2013). Statistical analysis of Hit/Miss data. *Materials Evaluation*, *71*(3), 322–329.

Kojima, M., Takahashi, H., & Kikura, H. (2019). Evaluation of capabilities on ultrasonic testing examiners using probability of defect detection and cumulative failure probability. *Journal of Advanced Maintenance*, *11*(2), 65–78.

Kurz, J. H., Jüngert, A., Dugan, S., & Dobmann, G. (2012, April 16-20). Probability of detection (POD) determination using ultrasound phased array for considering NDT in probabilistic damage assessments. In *South-African Insitute for Non-destructive Testing: World Conference on Nondestructive Testing* (pp. 1-10). Durban, South Africa.

Kurz, J. H., Jüngert, A., Dugan, S., Dobmann, G., & Boller, C. (2013). Reliability considerations of NDT by probability of detection (POD) determination using ultrasound phased array. *Engineering Failure Analysis, 35*, 609-617. https://doi.org/10.1016/j.engfailanal.2013.06.008

Le Gratiet, L., Iooss, B., Blatman, G., Browne, T., Cordeiro, S., & Goursaud, B. (2017). Model assisted probability of detection curves: New statistical tools and progressive methodology. *Journal of Nondestructive Evaluation*, *36*(8), 1–12. https://doi.org/10.1007/s10921-016-0387-z

Lei, X., Wirdelius, H., & Rosell, A. (2022). Simulation-based investigation of a probability of detection (POD) model using phased array ultrasonic testing (PAUT) technique. *Journal of Nondestructive Evaluation*, *41*(2), 1–13. https://doi.org/10.1007/s10921-022-00873-2

Lembersky, L., Adams, R., Tamutus, T., & Watson, J. (2012). Suspension cable wire break monitoring using acoustic emission for economic and probability of detection advantages. In *Structural Materials Technology 2012* (pp. 169-176). PubGenius Inc.

Marcotte, O., & Liyanage, T. (2017). *Nondestructive examination (NDE) used fuel containers probability of detection for increased probability of detection*. The American Society for Nondestructive Testing.

Meeker, W. Q. (2000). A methodology for predicting probability of detection for ultrasonic testing. In *AIP Conference Proceedings* (Vol. 557, No. 1, pp. 1972-1978). AIP Publishing. https://doi.org/10.1063/1.1373994

Mohr, G. A., & Willems, P. (2008, October 25-28). Factors affecting Probability of Detection with Computed Radiography. In *17th World Conference on Nondestructive Testing* (pp. 1-8). Shanghai, China.

Peeters, J., Ibarra-Castanedo, C., Khodayar, F., Mokhtari, Y., Sfarra, S., Zhang, H., Maldague, X., Dirckx, J. J. J., & Steenackers, G. (2018). Optimised dynamic line scan thermographic detection of CFRP inserts using FE updating and POD analysis. *NDT and E International*, *93*, 141–149. https://doi.org/10.1016/j.ndteint.2017.10.006

Poudel, A., Galvan-Nunez, S., Lindeman, B., & Gonzalez, F. (2022). *A quantitative assessment of historical nondestructive evaluation (NDE) probability of detection (POD) data for railroad tank cars*. The American Society for Nondestructive Testing.

Rentala, V. K., Mylavarapu, P., & Gautam, J. P. (2018). Issues in estimating probability of detection of NDT techniques – A model assisted approach. *Ultrasonics*, *87*, 59–70. https://doi.org/10.1016/j.ultras.2018.02.012

Rentala, V. K., Mylavarapu, P., K.Gopinath, Gautam, J. P., & Kumar, V. (2016, November 3-5). Model assisted probability of detection for lognormally distributed defects. In *8th International Symposium on NDT in Aerospace* (pp. 1-8). Bangalore, India.

Ribay, G., Mahaut, S., Cattiaux, G., & Sollier, T. (2017, September 4-7). Assessment of the reliability of phased array NDT of coarse grain component based on simulation. In *Proceedings of the 7th European-American Workshop on Reliability of NDE* (pp. 1–8). Potsdam, Germany. http://www.nde-reliability.de/portals/nde17/BB/21.pdf

Rodat, D., Guibert, F., Dominguez, N., & Calmon, P. (2017). Operational NDT simulator, towards human factors integration in simulated probability of detection. In *AIP Conference Proceedings* (Vol. 1806, No. 1). AIP Publishing. https://doi.org/10.1063/1.4974719

Schneider, C. R. A., Sanderson, R. M., Carpentier, C., Zhao, L., & Nageswaran, C. (2012, September 11-13). Estimation of probability of detection curves based on theoretical simulation of the inspection process. In *51st Annual Conference of the British Institute of Non-Destructive Testing* (pp. 393–404). Northamptonshire, UK.

Spies, M., & Rieder, H. (2018, June 11-15). An approach to the question 'How to account for human error in MAPOD?' In *12th European Conference on Non-Destructive Testing (ECNDT 2018)* (pp. 1–5). Gothenburg, Sweden. https://www.ndt.net/article/ecndt2018/papers/ecndt-0571-2018.pdf

Stubbs, D. A. (2005). Probability of detection for embedded defects: Needs for ultrasonic inspection of aerospace turbine engine components. *AIP Conference Proceedings*, *760*, 1909–1916. https://doi.org/10.1063/1.1916903

Subair, S. M., Balasubramaniam, K., Rajagopal, P., Kumar, A., Rao, B. P., & Jayakumar, T. (2014). Finite element simulations to predict probability of detection (PoD) curves for ultrasonic inspection of nuclear components. *Procedia Engineering*, *86*, 461–468. https://doi.org/10.1016/j.proeng.2014.11.059

Tai, J. L., Grzejda, R., Sultan, M. T. H., Łukaszewicz, A., Shahar, F. S., Tarasiuk, W., & Rychlik, A. (2023). Experimental investigation on the corrosion detectability of A36 low carbon steel by the method of phased array corrosion mapping. *Materials*, *16*(15), Article 5297. https://doi.org/10.3390/ma16155297

Tisseur, D., Costin, M., Fournier, S., Reece, C., & Schumm, A. (2019, October 1-3). POD calculation on a radiographic weld inspection with CIVA 11 RT module. In *10th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurized Components (JRC-NDE 2013)* (pp. 123–129). Cannes, France.

Topp, M., & Strothmann, L. (2021). How can NDT 4.0 improve the Probability of Detection (POD)? *E-Journal of Nondestructive Testing (NDT)*, *26*(4), 1–10. http://www.ndt.net/?id=26013

Tschoke, K., Mueller, I., Memmolo, V., Moix-Bonet, M., Moll, J., Lugovtsova, Y., Golub, M., Venkat, R. S., & Schubert, L. (2021). Feasibility of model-assisted probability of detection principles for structural health monitoring systems based on guided waves for fiber-reinforced composites. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, *68*(10), 3156–3173. https://doi.org/10.1109/TUFFC.2021.3084898

Underhill, P. R., & Krause, T. W. (2011). Quantitative fractography for improved probability of detection (POD) analysis of bolt hole eddy current. *Research in Nondestructive Evaluation*, *22*(2), 92–104. https://doi.org/10.1080/09349847.2011.553349

Underhill, P. R., & Krause, T. W. (2016). Eddy current probability of detection for mid-bore and corner cracks in bolt holes of service material. *Research in Nondestructive Evaluation*, *27*(1), 34–47. https://doi.org/10.1080/09349847.2015.1045642

Underhill, P. R., Uemura, C., & Krause, T. W. (2018). Probability of detection for bolt hole eddy current in extracted from service aircraft wing structures. In *AIP Conference Proceedings* (Vol. 1949, No. 1). AIP Publishing. https://doi.org/10.1063/1.5031619

Virkkunen, I., Koskinen, T., Papula, S., Sarikka, T., & Hänninen, H. (2019). Comparison of â versus a and Hit/Miss POD-estimation methods: A European viewpoint. *Journal of Nondestructive Evaluation*, *38*, 1-13. https://doi.org/10.1007/s10921-019-0628-z

Volker, A. W. F., Dijkstra, F. H., Terpstra, S., Heerings, H. A. M., & Lont, M. A. (2004, August 30 – September 3). Modeling of NDE reliability: Development of a “POD-Generator”. In *Proceedings of the 16th World Conference on Nondestructive Testing* (pp. 1-8). Montreal, Canada.

Wall, M., Burch, S., & Lilley, J. (2009). Human factors in POD modelling and use of trial data. *Insight: Non-Destructive Testing and Condition Monitoring*, *51*(10), 553–561. https://doi.org/10.1784/insi.2009.51.10.553

Wright, M. (2016, November 15-17). How to implement a PoD into a highly effective inspection strategy. In *NDT in Canada 2016 & 6th International CANDU In-Service Inspection Workshop* (pp. 1-8). Burlington, Canada. https://www.ndt.net/search/docs.php3?id=20418

Xu, Z., Zhou, Z., Chen, H., Qu, Z., & Liu, J. (2023). Effects of the wire mesh on pulsed eddy current detection of corrosion under insulation. *Nondestructive Testing and Evaluation*, *38*(2), 233–253. https://doi.org/10.1080/10589759.2022.2102167

Yosifov, M., Reiter, M., Heupl, S., Gusenbauer, C., Fröhler, B., Fernández- Gutiérrez, R., De Beenhouwer, J., Sijbers, J., Kastner, J., & Heinzl, C. (2022). Probability of detection applied to X-ray inspection using numerical simulations. *Nondestructive Testing and Evaluation*, *37*(5), 536–551. https://doi.org/10.1080/10589759.2022.2071892

Yosifov, M., Weinberger, P., Reiter, M., Fröhler, B., Beenhouwer, J. De, Sijbers, J., Kastner, J., & Heinzl, C. (2023). Defect detectability analysis via Probability of defect detection between traditional and deep learning methods in numerical simulations. *E-Journal of Nondestructive Testing*, *28*(3), 2–11. https://doi.org/10.58286/27716

Yusa, N. (2017). Probability of detection model for the non-destructive inspection of steam generator tubes of PWRs. *Journal of Physics: Conference Series*, *860*(1), 6–13. https://doi.org/10.1088/1742-6596/860/1/012032

Yusa, N., Chen, W., & Hashizume, H. (2016). Demonstration of probability of detection taking consideration of both the length and the depth of a flaw explicitly. *NDT and E International*, *81*, 1–8. https://doi.org/10.1016/j.ndteint.2016.03.001

Yusa, N., Tomizawa, T., Song, H., & Hashizume, H. (2018). Probability of detection analyses of eddy current data for the detection of corrosion. *Nondestructive Testing and Diagnostics*, *4*, 3–7. https://doi.org/10.26357/BNiD.2018.031

Zhao, J., Yang, K., Du, X., Yao, S., & Zhao, Y. (2023). Automated quantification of small defects in ultrasonic phased array imaging using AWGA-gcForest algorithm. *Nondestructive Testing and Evaluation*, 1–22. https://doi.org/10.1080/10589759.2023.2274001

Zhu, J., Min, Q., Wu, J., & Tian, G. Y. (2018). Probability of detection for eddy current pulsed thermography of angular defect quantification. *IEEE Transactions on Industrial Informatics*, *14*(12), 5658–5666. https://doi.org/10.1109/TII.2018.2866443