Pertanika J. Sci. & Technol. 33 (4): 1887 - 1906 (2025)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

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

Analysis of Phased Array Corrosion Mapping Data Using Probabilistic Detection (POD) Method

Jan Lean Tai[1], Mohamed Thariq Hameed Sultan[1, 2, 3]* and Farah Syazwani Shahar[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 Forestry and Forest Products (INTROP), Universiti 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

Phased Array Corrosion Mapping (PACM) has emerged as a powerful tool for detecting localized corrosion in industrial settings, particularly in the petrochemical industry. However, its application has been limited to plant operating temperatures within the recommended range for on-stream ultrasonic testing. This study addresses this gap by conducting a comprehensive Probability of Detection (POD) analysis using the MH1823 POD software, following the guidelines of MIL-HDBK-1823A (Department of Defense Handbook, 2009). The binary hit/miss method was employed to evaluate the accuracy of PACM on carbon steel (CS) and stainless steel (SS 304 and SS 316) surfaces at elevated temperatures up to 250 °C. Experimental results demonstrated successful detection of all flat-bottom-hole (FBH) indications, with deviations categorized as "miss" when exceeding 10% of the designed dimension. The POD curves revealed robust detection capabilities across various temperatures, with distinct clusters exhibiting different a50, a90, and a90/95 values. These findings underscore the reliability and efficacy of PACM in high-temperature applications, offering a significant advancement in non-destructive testing (NDT) for the petrochemical industry. The study highlights the potential of PACM to enhance turnaround maintenance (TAM) by enabling accurate on-stream inspections, thereby reducing downtime and increasing operational efficiency. Future research should explore the "â versus a" method to further enhance the accuracy and understanding of PACM in NDT applications.

Keywords: Non-destructive testing, phased array ultrasonic testing, probability of detection, turnaround maintenance

ARTICLE INFO

Article history:
Received: 19 December 2024
Accepted: 27 March 2025
Published: 11 June 2025

DOI: https://doi.org/10.47836/pjst.33.4.10

E-mail addresses:
taijanlean2008@hotmail.com (Jan Lean Tai)
thariq@upm.edu.my (Mohamed Thariq Hameed Sultan)
farahsyazwani@upm.edu.my (Farah Syazwani Shahar)
* Corresponding author

e-ISSN: 2231-8526

INTRODUCTION

The continuous growth of the petrochemical industry necessitates the ongoing modification and expansion of existing factories to accommodate the production of additional by-products. However, these expansions result in increased complexity and maintenance requirements. Consequently, many petrochemical plants undergo periodic shutdowns and turnaround maintenance (TAM) after several years of operation. TAM is a critical process for addressing various maintenance needs, primarily due to the aging of equipment and infrastructure. Over time, aging can lead to wear and tear, corrosion, and other forms of degradation, resulting in decreased efficiency, increased safety risks, and the potential for unplanned shutdowns or catastrophic failures (Aiello et al., 2020). Additionally, the buildup of fouling or contaminants in equipment can reduce efficiency and increase the risk of corrosion or other damage, necessitating TAM to address these issues (Al-Turki et al., 2019).

The demand for TAM arises from potential disruptions; even minor repairs can lead to economic losses due to halted production of multiple products. While all facility owners strive to avoid maintenance-related shutdowns, they are acutely aware of the risks posed by neglecting proper maintenance, which can lead to unplanned shutdowns and industrial accidents (Hlophe & Visser, 2018).

In response, many owners are exploring optimizing the TAM process by minimizing turnaround time or extending the intervals between TAM events to maximize operational benefits (Al-Marri et al., 2020; Elwerfalli et al., 2019). The petrochemical industry faces significant corrosion-related challenges, leading to decreased efficiency, increased safety risks, and potential unplanned shutdowns or catastrophic failures.

Ultrasonic thickness gauging (UTG) is a popular method for detecting corrosion during on-stream inspections in the petrochemical industry, primarily due to its cost-effectiveness and ease of use. Inspectors can quickly learn to operate UTG with minimal training. However, UTG has some limitations; its point-by-point nature can lead to inadequate detection of localized corrosion and make inspections of larger areas time-consuming.

In contrast, phased array corrosion mapping (PACM) offers a more comprehensive solution. It is important to note that, despite its advantages, there are currently no established examples of ultrasonic applications for high-temperature operations. The recommended temperature range for ultrasonic testing (UT) is up to 52 °C (Jory, 2019). However, on-stream inspections may require work beyond these limits.

Previous experiments have been conducted to investigate the feasibility of extending the use of UT beyond its recommended temperature limit (Tai, Sultan et al., 2023). The results indicate that as temperature increases, the velocity of ultrasound waves decreases while attenuation increases. The following research aimed to build upon the previous experiment's findings and explore the effectiveness of PACM in detecting corrosion at

high temperatures. The results confirmed that PACM could accurately detect corrosion in high-temperature environments when the velocity and decibel settings were appropriately adjusted (Tai, Grzejda et al., 2023).

The utilization of Probability of Detection (POD) in Non-Destructive Testing (NDT) is of significant importance. POD provides a quantifiable measure of the success rate in identifying specific targets or flaws through designated inspection methods, thereby playing a crucial role in evaluating the effectiveness of novel NDT techniques (Cherry & Knott, 2022).

A review of academic articles published over the past two decades focusing on POD in relation to NDT reveals that UT and Eddy Current Testing (ECT) were among the initial methods to employ POD for assessing detection probabilities (Bato et al., 2020; Goursolle et al., 2016; Yusa, 2017). Following the development of POD and the introduction of Model-Assisted Probability of Detection (MAPOD) (Rodat et al., 2017), the application of POD has significantly expanded to encompass various NDT methods.

The literature review initially aimed to identify prior instances of POD applications in PACM. However, the limited availability of literature on this subject prompted a shift in focus toward using Phased Array Ultrasonic Testing (PAUT) (Bajgholi et al., 2023). While these two inspection methods differ in orientation, they share several commonalities.

A consistent factor contributing to the success of the output results is the incorporation of uncertainty parameters, as highlighted by numerous scholars in the literature. The selection of these parameters, including defect size, material thickness (Ribay et al., 2017), and water path considerations, plays a crucial role in PAUT, similar to conventional UT. Additionally, both methods utilize calibration blocks featuring flat bottom holes (FBH) and side drill holes (SDH) to calibrate the equipment (Dominguez et al., 2016; Marcotte & Liyanage, 2017).

This study aims to assess the accuracy of novel PACM testing techniques on surfaces with elevated temperatures of up to 250 °C through POD analysis. Upon reviewing the previously mentioned articles, it was noted that despite applying POD analyses to numerous NDT methods, such analyses were significantly absent for PACM. This gap represents the distinctive novelty of the current study.

METHODOLOGY

There are two distinct methods for conducting POD analysis: the “hit/miss” method and the “signal response versus flaw size” (â versus a) method. The choice between these methods depends on the nature of the data generated by the NDT system under examination (Rentala et al., 2018).

The “hit/miss” method is applied when the NDT system produces binary outcomes, meaning a flaw is either detected or not detected. In this method, the POD curve is estimated based on the proportion of detected flaws to the total number of flaws present in the sample

(Knott & Kabban, 2022b). This approach is typically used when the inspection data is discrete and lacks a measurable response associated with flaw detection. For example, in PT or RT, a flaw is either observed or not, with no measurable response linked to the detection.

In contrast, the "â versus a" method is utilized when the NDT system yields measurable responses that can be correlated with the size of the flaw. This method establishes a relationship between the signal response (â) and the flaw size (a), enabling the estimation of the POD curve based on the distribution of signal responses for a given flaw size. This approach is commonly used in NDT methods that provide measurable responses, such as UT and Eddy ECT (Kojima et al., 2019; Yusa et al., 2018; Zhu et al., 2018). It allows for assessing the system's sensitivity to different flaw sizes and provides a quantitative measure of its performance.

The "hit/miss" approach relies on examining the number of successful detections (hits) and false negatives (misses) in a series of inspections. Logistic regression is a widely used statistical method in this approach, establishing a model for the probability of detection as a function of various factors, such as flaw size, inspection conditions, and system parameters (Knott & Kabban, 2022a).

In the "â versus a" approach, regression analysis with censored data is the primary statistical technique. This method is employed when some flaw sizes remain unidentified (censored) by the NDT system. The goal is to address the censoring of non-detected flaws and estimate the probability distribution of the detected flaw sizes. The regression model can then be used to determine the mean and variance of the detected flaw size distribution, allowing for calculating the probability of detection curve based on the estimated distribution (Rentala & Kanzler, 2022).

The accuracy of the output is contingent upon the amount of input data; insufficient data can introduce bias. MIL-HDBK-1823A recommends analyzing a minimum of 40 representative defect data points for the "â versus a" method, alongside signal strength ("â") measurements and crack sizes ("a") (Department of Defense Handbook, 2009). In contrast, the "hit/miss" method does not process signal values. Instead, it estimates the POD curve based on binary results, where a hit indicates that a crack was detected and a miss signifies that it was not. This method necessitates more data due to the limited information it yields, requiring at least 60 data sets to avoid bias and ensure a reliable POD curve (Virkkunen et al., 2019).

The a90/95 value derived from the estimated POD curve is particularly significant in these applications. This value indicates the defect size that can be detected with a 90% probability and a 95% confidence interval, thereby providing a tangible measure of the reliability and effectiveness of the NDT process (Kim et al., 2021).

In a conventional two-parameter model, three key metrics are of interest: **a50**, **a90**, and **a90/95**. These parameters reflect different aspects of the Probability of Detection

(POD) curve:

- **a50**: This represents the discontinuity size for which the mean POD is 50%. It indicates the size at which the probability of detection is statistically expected to be 50%.
- **a90**: This is the discontinuity size for which the mean POD is 90%. It denotes the size at which the probability of detection is expected to be approximately 90%.
- **a90/95**: This parameter refers to the upper 95% confidence bound on a90, providing a level of certainty that the true a90 lies below this value, with 95% confidence.

The two-parameter model imposes a specific behavior on the POD curve: As the size of the discontinuity approaches zero, the POD should converge to zero; conversely, as the size approaches infinity, the POD should approach one (Knopp et al., 2019).

Binary Hit/Miss Approach

Several possible outcomes can result from NDT evaluations aimed at identifying defects. These outcomes include:

1. **True Positive (TP)**: A defect exists and is correctly detected by the assessment.
2. **False Positive (FP)**: No defects exist, but an incorrect assessment indicates the presence of defects.
3. **False Negative (FN)**: A defect exists, but the assessment fails to detect it.
4. **True Negative (TN)**: No defect exists, and the assessment correctly indicates the absence of a defect.

Two critical, independent probabilities are considered to quantify the capability of the NDT methods (Keprate & Ratnayake, 2015):

- **Probability of Detection (POD)** or probability of true positives (P(TP))
  This represents the likelihood that the assessment will correctly detect a defect. It is calculated using the following Equation 1:

$$POD\ or\ P(TP) = \frac{TP}{TP+FN} \qquad [1]$$

- **Probability of False Alarm (POFA)** or probability of false positives (P(FP))
  This represents the probability that the assessment incorrectly indicates the presence or absence of a defect. It is calculated using the following Equation 2:

$$POFA\ or\ P(FP) = \frac{FP}{TP+FP} \qquad [2]$$

Signal Response "â versus a" Approach

The "â versus a" method is utilized in NDT when measurable responses, typically in the form of signal amplitudes, are generated by the NDT system instead of binary outcomes

like “hit” or “miss.” This approach is particularly useful for NDT systems that produce quantitative data, allowing for a more comprehensive system performance evaluation. It is commonly employed in techniques such as UT and ECT, where signal amplitude provides valuable information about the size and characteristics of defects.

In the “â versus a” method, specifically in UT, the echo amplitude “â” is proportional to the measured crack size “a.” This relationship helps determine if the echo amplitude “â” is degraded due to noise. Given the variability in the relationship between “â” and “a,” the goal is to establish a threshold value for “â” that maximizes crack detection while minimizing false alarms caused by noise. To address this challenge, ASTM E3023 and MIL-HDBK-1823A utilize a linear function to relate “â” to “a” and calculate prediction intervals to account for noise and statistical uncertainty (ASTM E3023-21, 2021; Department of Defense Handbook, 2009).

This method allows for assessing the POD of a flaw based on the measured signal response and the flaw’s characteristics. However, unlike the hit/miss method, it lacks a single, straightforward formula.

PACM Experiment and Data Acquisition

PACM and PAUT share many similarities in the processes and equipment used, with the primary distinction in their defect detection approaches. PAUT employs steering and focusing of ultrasound beams to identify weld defects, whereas PACM maintains a normal beam orientation to detect material corrosion, lamination, or metal loss. This wave propagation resembles that of a 0-degree normal probe in UT. Both PAUT and PACM utilize A-, B-, C-, and S-scan displays to provide a comprehensive three-dimensional representation of inspection results, as illustrated in Figure 1.

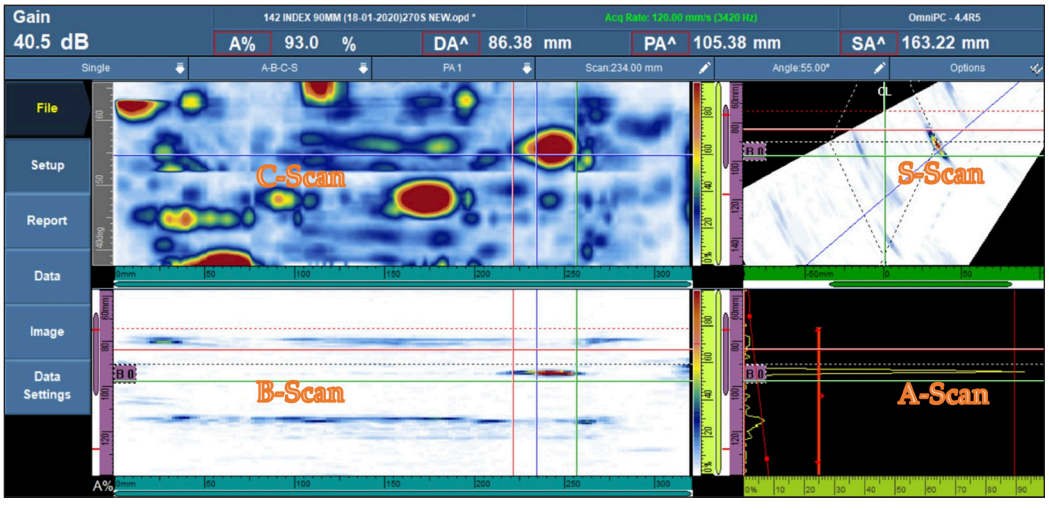

Figure 1. PAUT and PACM result presentation

The A-Scan displays the reflected signals and measures ultrasound echo amplitudes, integral to the quantitative measurement of the "â versus a" method used in POD analysis. In contrast, the B-Scan (side view), S-Scan (sectoral view), and C-Scan (top view) provide qualitative defect information relevant to the "hit/miss" method for defect detection.

Selecting appropriate calibration standards is critical in PACM experiments. Choosing between side drill holes (SDH) and FBH is particularly important. While SDHs are used in some NDT applications, they may not be suitable for PACM calibration, especially with a normal beam ultrasound configuration, where the ultrasound travels perpendicular to the test piece surface.

The test blocks designed for the PACM experiment comprised FBHs of various diameters and depths, as well as surface slots with varying depths, as shown in Figure 2. This design was compared to three different materials: (1) CS, (2) stainless steel 304 (SS 304), and (3) stainless steel 316 (SS 316).

The PACM experiment adhered to a standardized protocol from previous studies aligned with industry guidelines. The protocol specifies calibration procedures, transducer selection, and data acquisition parameters to ensure the consistency and comparability of the test results (Tai, Grzejda et al., 2023; Tai, Sultan, Shahar, Łukaszewicz et al., 2024).

A test specimen was subjected to controlled heating using a heat treatment machine, incrementally raising the temperature from 30 °C to 250 °C in 10-degree increments. This temperature variation aimed to mimic real-world conditions and assess the performance of the PACM under elevated temperature scenarios.

In the experimental setup, the specimens used for data collection included FBHs labeled A1 to A4 and B1 to B4, each with receptivity depths of 3 mm and 6 mm. Additionally, surface notches labeled C1 to C4 were incorporated, all having a consistent width of 5 mm but varying depths. As illustrated in Figure 3, which presents the C-scan results of the PACM experiment, this technique provides accurate depth measurements and visually differentiates depths using color coding. Furthermore, the size of the FBHs can be quantified using specialized software.

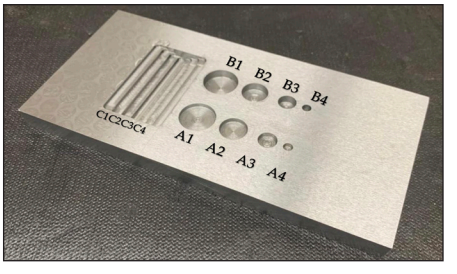

Figure 2. Schematic of the PACM experiment specimen design

POD Method with Mh1823 Software

The MIL-HDBK-1823A manual provides a comprehensive guide for constructing POD studies (Department of Defense Handbook, 2009). It also includes a useful guide for downloading Mh1823 POD software, which is built on the R statistical and graphics engine (https://www.r-project.org/). This software is accessible through the Statistical Engineering website developed by Annis (2023).

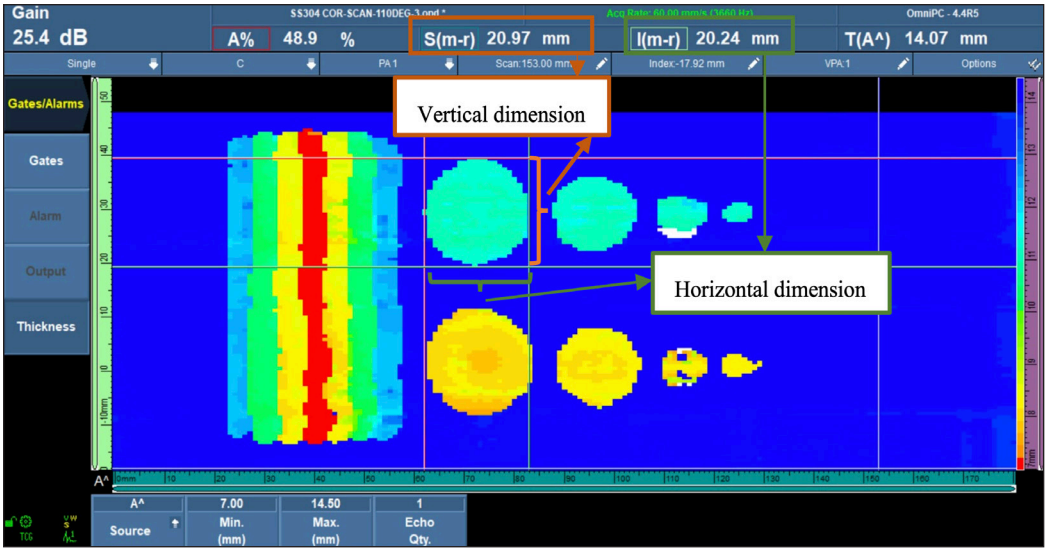

Figure 3. PACM data example to indicate the vertical and horizontal dimensions

The binary hit/miss POD approach was selected for analyzing PACM results due to its clear graphical representation and ability to measure flaw sizes. The following definitions were used for POD curve generation:

- The determination of hit/miss status depended on the precision of the results obtained, with all indications considered significant.
- Depth indications and dimensions of notches C1 to C4 were excluded because of consistent deviations below 0.5 mm.
- Measurements of specimens FBH A1–A4 and B1–B4, each with varying depths, were taken both vertically (VD) and horizontally (HD) to assess their variability concerning specimen dimensions (SD).
- A total of 138 deviation data points were collected for each FBH, resulting in an aggregate of 1,104 data points across all eight FBH indications. Example experimental data for FBH A1 are presented in Table 1.
- Instances of deviation equal to or exceeding 10% of the designated size were classified as “miss,” while those below this threshold were labeled “hit.”
- The POD curve was generated based on the same FBH dimension data for each temperature across the three different material types.

Table 1 demonstrates a systematic deviation increase for CS test results as temperature rises, particularly above 150 °C. Key observations include:

- At **30 °C**, deviations for all FBHs (A1–A4, B1–B4) remain ≤1 mm.
- At **190 °C**, vertical deviations (VD) reach 1.2 mm (FBH A1), and horizontal deviations (HD) peak at 0.79 mm.

Table 1
Experiment PACM data for carbon steel FBH A1

	CS test 1					CS test 2				CS test 3			
	SD	**VD**	**D**	**HD**	**D**	**VD**	**D**	**HD**	**D**	**VD**	**D**	**HD**	**D**
30 °C	20	21	1	20.01	0.01	20.7	0.7	19.69	0.31	19.31	0.69	20.06	0.06
40 °C	20	20.7	0.7	20.17	0.17	20.1	0.1	19.85	0.15	20.25	0.25	20.48	0.48
50 °C	20	19.7	0.3	19.85	0.15	20	0	19.69	0.31	21.6	1.6	20.32	0.32
60 °C	20	19.5	0.5	20.01	0.01	20.2	0.2	20.48	0.48	19.53	0.47	20.01	0.01
70 °C	20	20.8	0.8	20.17	0.17	21.2	1.2	20.05	0.05	21.32	1.32	20.32	0.32
80 °C	20	20.8	0.8	19.91	0.09	21.5	1.5	20.64	0.64	19.5	0.5	20.48	0.48
90 °C	20	21.1	1.1	19.63	0.37	21	1	20.52	0.52	20.24	0.24	20.35	0.35
100 °C	20	20.3	0.3	20.05	0.05	20.5	0.5	19.05	0.95	19.8	0.2	19.37	0.63
110 °C	20	20.5	0.5	20.15	0.15	20	0	19.21	0.79	20.42	0.42	19.15	0.85
120 °C	20	20.6	0.6	19.59	0.41	19.71	0.29	19.37	0.63	20.07	0.07	19.37	0.63
130 °C	20	20.4	0.4	19.53	0.47	20.43	0.43	19.15	0.85	20.43	0.43	19.04	0.96
140 °C	20	19.5	0.5	19.63	0.37	20.25	0.25	20.48	0.48	20.25	0.25	20.48	0.48
150 °C	20	19.6	0.4	19.53	0.47	20.45	0.45	20.32	0.32	19.89	0.11	20.64	0.64
160 °C	20	20.5	0.5	20.01	0.01	20.61	0.61	19.47	0.53	20.25	0.25	20.01	0.01
170 °C	20	20.7	0.7	19.69	0.31	21.17	1.17	20.17	0.17	21.32	1.32	20.45	0.45
180 °C	20	21	1	20.03	0.03	21.14	1.14	20.17	0.17	21.14	1.14	21.12	1.12
190 °C	20	21.2	1.2	19.21	0.79	20.07	0.07	20.17	0.17	20.79	0.79	20.76	0.76
200 °C	20	21.3	1.3	19.15	0.85	20.43	0.43	19.53	0.47	20.79	0.79	20.7	0.7
210 °C	20	21.1	1.1	20.21	0.21	21.14	1.14	20.48	0.48	20.07	0.07	20.01	0.01
220 °C	20	20.5	0.5	20.17	0.17	19.95	0.05	20.64	0.64	19.71	0.29	19.05	0.95
230 °C	20	21	1	20.41	0.41	20.78	0.78	20.64	0.64	19.8	0.2	20.05	0.05
240 °C	20	20	0	20.17	0.17	20.96	0.96	20.96	0.96	20.61	0.61	19.53	0.47
250 °C	20	20.2	0.2	20.64	0.64	20.79	0.79	20.7	0.7	19.53	0.47	20.64	0.64

Notes. All units in mm. SD = specimen diameter;VD = vertical dimension; HD = horizontal dimension; D = deviation

- At **250 °C**, HD deviations for FBH B1 exceed 0.7 mm, approaching the 10% threshold for "miss" classification.

RESULTS AND DISCUSSION

The "hit/miss" model is a core component of POD analysis, facilitating the quantification of the likelihood of correctly detecting flaws or defects. This model offers four distinct link functions, each with unique mathematical characteristics:

1. Logit, Logistic, or Log-Odds Function: The Logit function is widely used in POD analysis due to its compatibility with binary outcomes. It transforms the probability of detection into log-odds, which aids in regression analysis and curve fitting.
2. Probit or Inverse Normal Function: Probit analysis serves as an alternative approach

that utilizes the cumulative distribution function of a standard normal distribution. This method can be advantageous when the logit function does not adequately fit the data.

3. Complementary Log-Log Function, Weibull: The Complementary Log-Log function effectively models rare events and tail probabilities, making it useful when dealing with extreme values during detection.
4. Loglog Function: The Loglog function is another option for modeling POD data, although it is less commonly utilized. This function is particularly beneficial for handling complex data distributions.

The MH1823 POD software allows users to select the most appropriate link function after inputting the data points. It provides eight possible POD curves, comprising four link functions for selection, as illustrated in Figure 4.

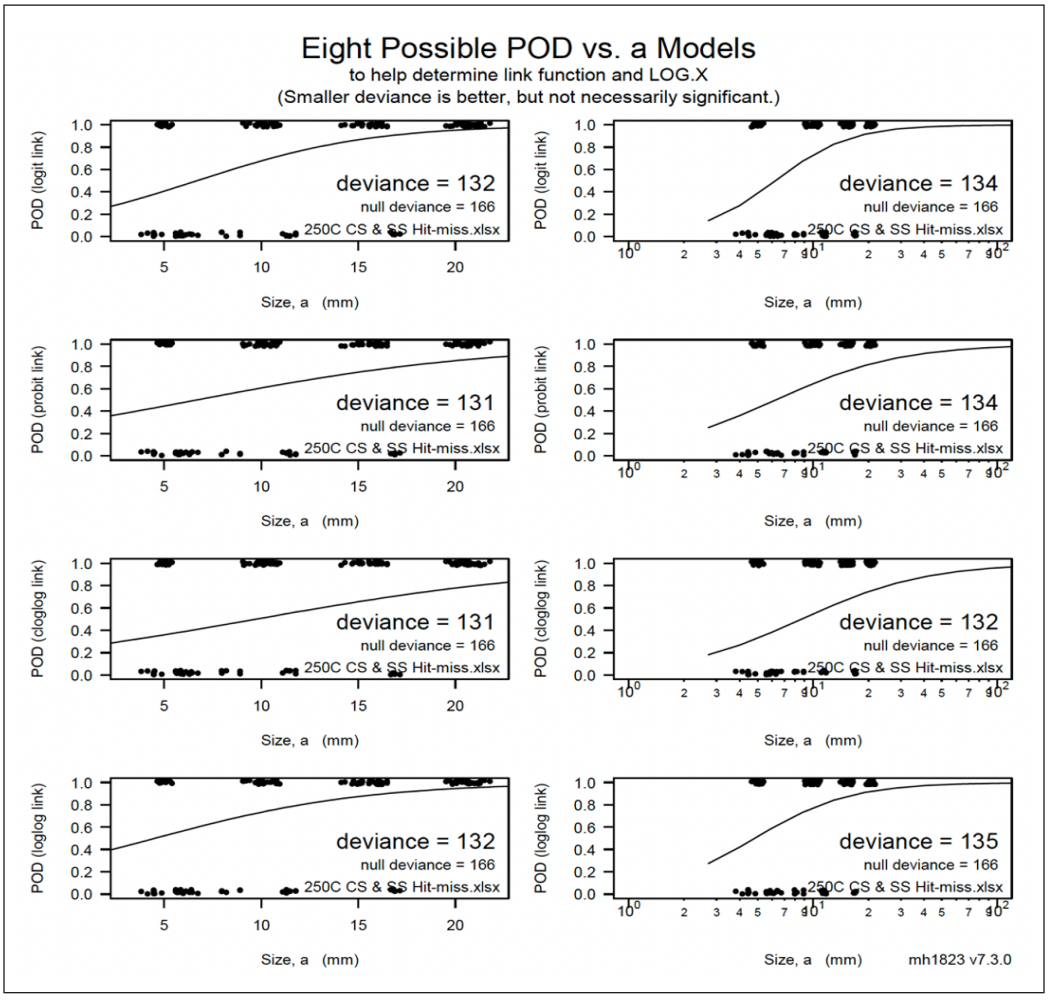

Figure 4. Eight possible POD curves generated from four-link functions

In this study, all POD curves are modeled using the logit function, which is defined as the natural logarithm of the odds ratio. The odds ratio represents the ratio of the POD to the probability of non-detection. Mathematically, the logit function is expressed as Equation 3.

$$Logit(P) = \ln(\frac{P}{1-P}) \quad [3]$$

Where P denotes the probability of detection, ln denotes the natural logarithm function.

Applying the logit function transforms the POD, originally ranging from 0 to 1, into a scale from negative to positive infinity. This transformation facilitates using linear regression techniques, which are simpler and more interpretable compared to non-linear models.

A notable advantage of the logit function is its symmetry around P=0.5. This symmetry implies that equal weight is assigned to both the probability of detection and non-detection, which is particularly beneficial in POD analysis, where both false positives and false negatives are significant.

The POD curve typically consists of a solid line and dotted lines. The solid line represents the estimated POD for a given flaw size, while the dotted lines indicate the confidence bounds around the estimated POD. These bounds reflect the level of uncertainty associated with the estimate, with their width depending on the amount and variability of the data. The solid line is the primary focus, providing an estimate of the probability of detecting a flaw of a given size. In contrast, the dotted lines convey the reliability of that estimate.

It is important to clarify that "a50" does not indicate the discontinuity size at which the mean POD reaches 50%, nor does it imply that 50% of the PACM indications align precisely with this size without deviation. Similarly, "a90" refers to the size at which the detection probability is expected to be approximately 90%.

A meticulous examination of the plotted POD curves reveals that PACM demonstrates robust detection capabilities across various materials and temperatures, ranging from 30 °C to 250 °C. The Hit/Miss POD results prioritize the accuracy of identified flaws, ensuring deviations remain below the 10% threshold.

This analysis underscores the reliability and efficacy of PACM in identifying flaws within diverse materials over a wide operational temperature range. The emphasis on minimizing deviations in the Hit/Miss POD approach further enhances the precision of defect identification and characterization.

Upon comparing all POD curves, four distinct clusters are observed. Table 2 presents the a50, a90, and a90/95 values for all temperatures from 30 °C to 250 °C across these clusters.

The variability in POD outcomes can be attributed to inherent differences in data combinations and collection methodologies. This variance highlights the absence of a consistent pattern in the POD curves, emphasizing the importance of considering these factors when interpreting the results.

The first cluster, illustrated in Figure 5, was formed at 190 °C. It displays a a50 of 6.655 mm, indicating that 50% of the FBHs larger than this size exhibit no deviation. In comparison, a90 at 17.67 mm denotes a 90% probability of no deviation for FBHs with dimensions equal to or greater than 17.67 mm. Additionally, the a90/95 value is 22.72 mm, which exceeds all processed data. The values of a50, a90, and a90/95 can be obtained from the graph (highlighted in red). This group predominantly includes temperatures of 60 °C, 70 °C, 100 °C, 110 °C, 130 °C, 140 °C, 150 °C, 160 °C, 190 °C, and 200 °C.

Table 2
a50, a90, and a90/95 in four clusters

	Temp (°C)	a50	a90	a90/95
Cluster 1	60	5.539	16.89	22.09
	70	7.372	17.98	22.84
	100	7.572	19.91	26.11
	110	7.213	18.27	23.36
	130	6.574	17.33	22.24
	140	3.844	17.53	24.65
	150	5.279	17.85	24.06
	160	5.459	17.16	22.67
	190	6.655	17.67	22.72
	200	8.521	18.27	22.59
Cluster 2	80	5.948	15.01	18.84
	90	6.239	16.46	20.96
	120	8.202	15.43	18.44
	170	4.041	16.02	21.62
	250	6.692	16.49	20.77
Cluster 3	30	2.153	22.46	NA
	40	1.811	23.77	NA
	50	0.576	22.70	NA
	210	1.684	22.25	NA
Cluster 4	220	6.523	22.41	32.77
	230	3.784	21.94	35.74
	240	6.555	21.98	31.68

The second cluster includes temperatures of 80 °C, 90 °C, 120 °C, 170 °C, and 250 °C, with one of the curves representing data collected at 250 °C, as shown in Figure 6. At this temperature, the recorded values for a50, a90, and a90/95 offer valuable insights into the flaw detection capabilities of the PACM technique. The a50 value at 250 °C is 6.692 mm, indicating a 50% probability of detecting a flaw with a diameter of 6.692 mm or larger. This suggests that the PACM technique maintains a relatively high level of accuracy in detecting flaws at this temperature, even for smaller flaw sizes.

The value of a90 at 250 °C is 16.49 mm, indicating a 90% probability of detecting a flaw with a diameter of 16.49 mm or larger. This suggests that the PACM technique is highly reliable in detecting larger flaws at high temperatures, with a low rate of false negatives. The value of a90/95 at 250 °C is 20.77 mm, which indicates a 95% probability of detecting a flaw with a diameter of 20.77 mm or larger, given that a flaw with a diameter of 16.49 mm or larger has already been detected with 90% probability. This value provides insight into the consistency and reliability of the PACM technique in detecting flaws at high temperatures.

In the third cluster, which includes temperatures of 30 °C, 40 °C, 50 °C, and 210 °C, as shown in Figure 7, a noteworthy observation is made. At 30 °C, there is a significant reduction in the value of a50, measured at 2.153 mm, compared to the other clusters.

The POD curve at 30 °C shows a significant reduction in the value of a50, indicating that the probability of detecting a 5 mm diameter flaw is lower at this temperature compared

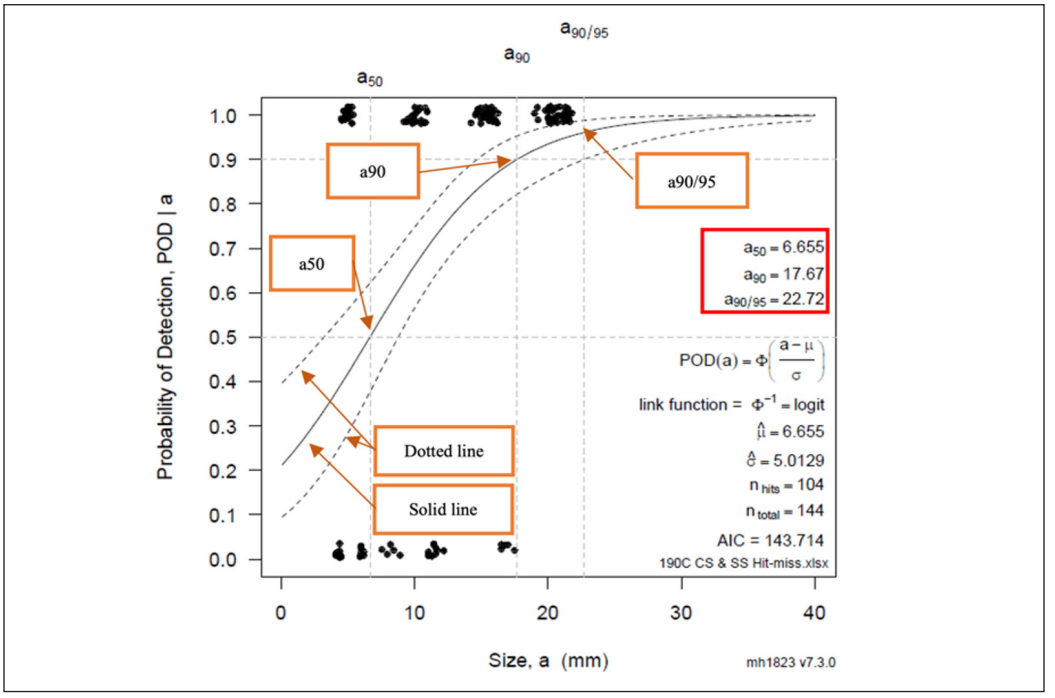

Figure 5. 190 °C POD curve

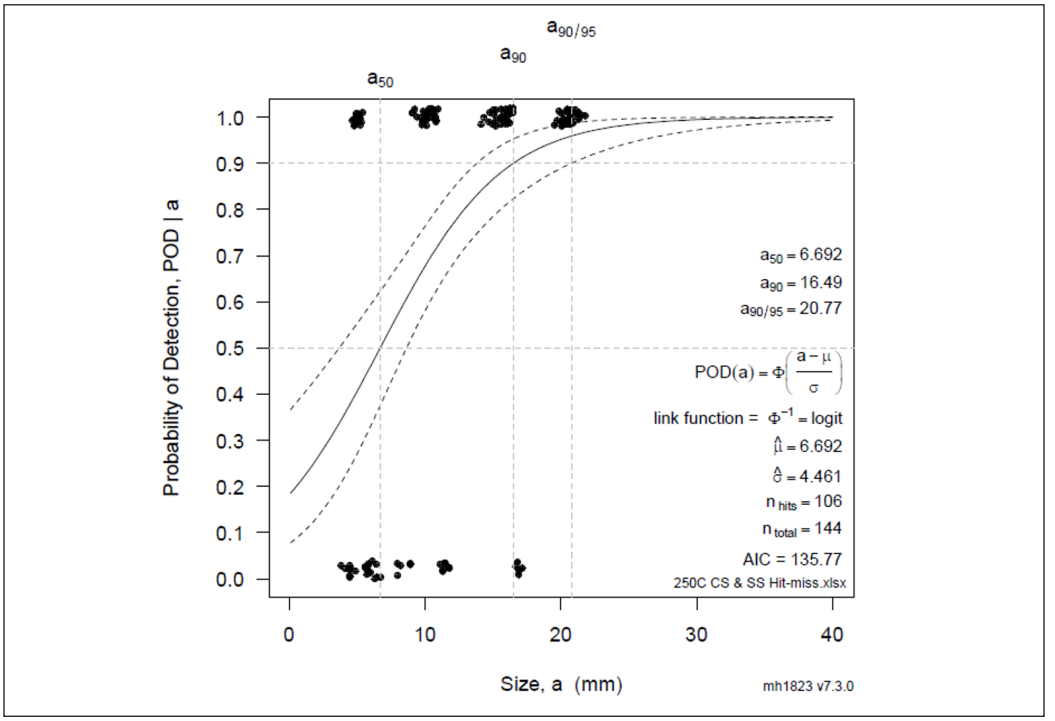

Figure 6. 250 °C POD curve

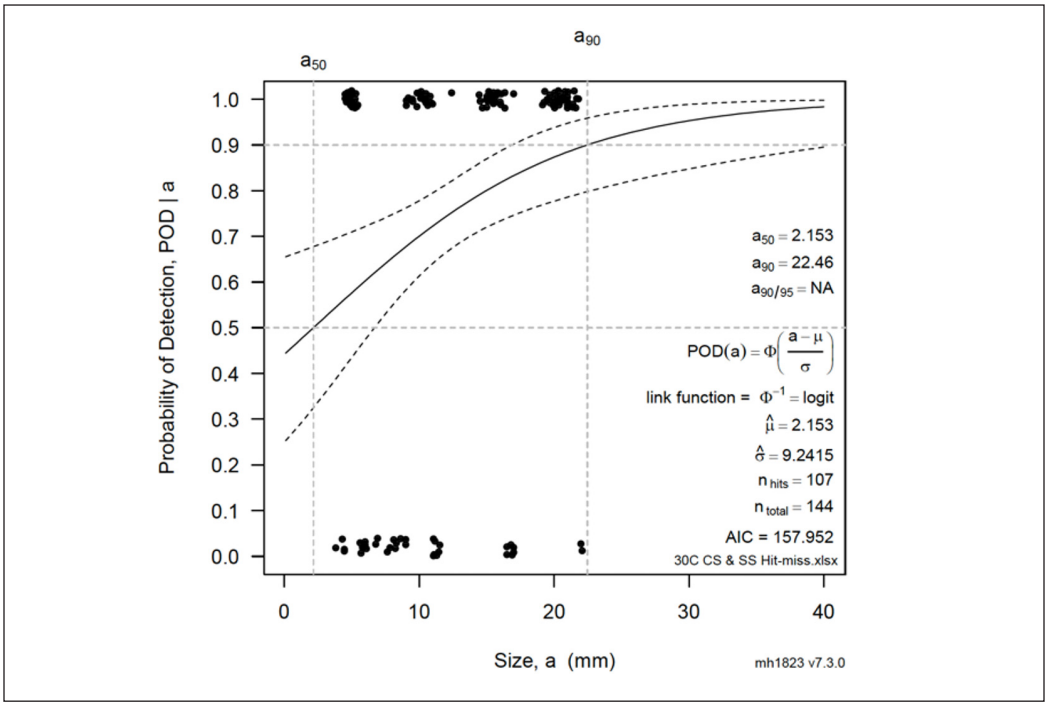

Figure 7. 30 °C POD curve

to others. This reduction can be attributed to the lower deviation observed in the 5 mm diameter FBH A4 and B4, which may be due to well-conditioned data or a high level of accuracy in the PACM technique at lower temperatures.

Furthermore, all graphs within this cluster indicate "NA" for a90/a95, meaning that the parameter does not exist at 30 °C. This occurs because the lower bound never reaches a POD value of 0.9, which is necessary to calculate a90/a95. This anomaly may also be attributed to well-conditioned data or a high level of accuracy in the PACM technique at lower temperatures, resulting in a POD ceiling where the maximum attainable POD is less than 1.

In summary, the lower deviation observed in the 5 mm diameter FBH A4 and B4, along with the absence of a90/a95 at 30 °C, suggests that the PACM technique demonstrates high accuracy without deviation in detecting smaller flaws at lower temperatures. However, data conditioning and the POD ceiling should still be considered when interpreting POD curves to ensure an accurate assessment of the PACM technique's performance, detecting flaws at different temperatures.

The fourth cluster, comprising temperatures of 220 °C, 230 °C, and 240 °C, exhibits distinct characteristics that set it apart from the other clusters. As depicted in Figure 8, the POD curve for this cluster at 230 °C shows a50 and a90 values of 3.784 mm and 21.94 mm, respectively. Notably, the hit-and-miss data patterns observed at this dimension resemble those in cluster 1.

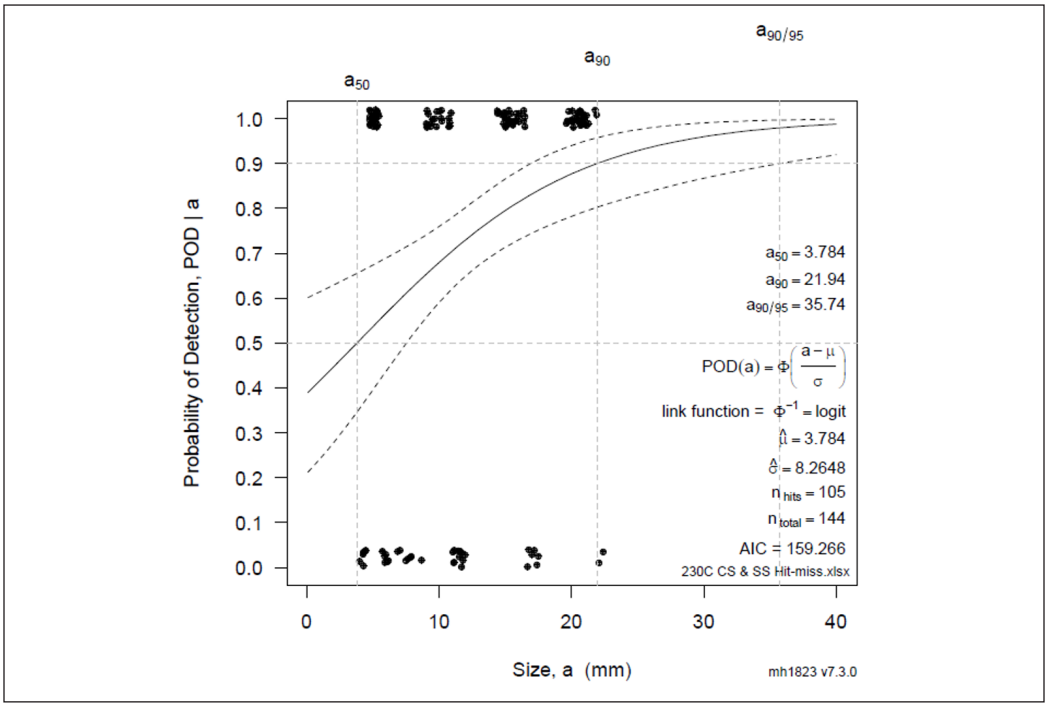

Figure 8. 230 °C POD curve

The above graphs demonstrate that most data have deviations below 10%. This indicates that PACM can be effectively utilized in high-temperature operations and across three types of materials.

POD Curve Analysis

The PACM technique demonstrated robust detection capabilities for larger flaws (≥17 mm) at all temperatures, with a90 values consistently exceeding 16 mm (Table 2). At high temperatures (e.g., 250 °C), the a90/95 value (20.77 mm) indicates a 95% confidence in detecting flaws ≥20.77 mm, which is comparable to low-temperature performance (e.g., 30 °C, a90 = 22.46 mm). However, smaller flaws (≤5 mm) showed reduced detectability at high temperatures due to increased ultrasonic attenuation and thermal noise. This suggests PACM is reliable for critical flaw detection in high-temperature environments but requires additional validation for minor defects.

Standardization of PACM Accuracy

To establish PACM accuracy, we propose adopting the MIL-HDBK-1823A guidelines for POD analysis, which define a90/95 as the critical metric for system reliability (Department of Defense Handbook, 2009; Tai, Sultan, Shahar, Yidris et al., 2024). Additionally,

referencing ASTM E2862-23 (2023) guidelines provides a standardized approach to POD analysis, critical for quantifying system reliability in NDT applications.

CONCLUSION

In the field of NDT, the POD serves as a critical metric for quantifying the likelihood of identifying specific targets or flaws through designated inspection methods. This study aimed to assess the effectiveness of PACM for detecting corrosion in high-temperature environments, specifically up to 250 °C, and to quantify its reliability through a POD analysis.

Our research has established that PACM can be effectively applied to CS, SS 304, and SS 316 surfaces within the experimental temperature range of 30 °C to 250 °C. Through comprehensive POD analysis using the binary hit/miss method, we have shown that PACM maintains robust detection capabilities across various materials and temperatures, with distinct clusters of a50, a90, and a90/95 values indicating consistent performance metrics.

The study contributes to the field by expanding the operational temperature range for reliable ultrasonic corrosion detection and offering a systematic approach to assessing PACM accuracy using POD analysis. This advancement addresses a significant gap in current industrial inspection capabilities, particularly for on-stream inspections in the petrochemical industry, where high-temperature operations are common.

While the hit/miss method provided valuable insights, future research should explore the "â versus a" method for POD analysis in PACM. This approach, which considers the size of detected flaws relative to their actual dimensions, has the potential to enhance the accuracy and depth of POD analysis. By incorporating A-Scan amplitude signals, this method could provide a more nuanced understanding of flaw-detection mechanisms and improve the precision of PACM inspections.

However, implementing the "â versus a" method presents practical challenges, particularly in data acquisition and comparison. The method requires extensive datasets to establish reliable relationships between detected and actual flaw sizes, which demands significant resources and time. Additionally, differences in data analysis techniques between the "â versus a" and hit/miss methods may complicate direct comparisons of results. Addressing these challenges will require standardized protocols and potentially new analytical frameworks to ensure consistent and comparable outcomes across different inspection scenarios.

The current study has limitations that should be acknowledged. Our research focused on three specific materials (CS, SS 304, SS 316) and relatively simple flaw geometries. While these materials are commonly used in industrial applications, the universal applicability of PACM needs further validation across diverse alloys and more complex defect configurations. Additionally, our laboratory-based experiments may not fully capture the variability and challenges of real-world inspection environments.

To build upon this research, we recommend several specific directions for future work:

1. Integrating the "â versus a" method into PACM POD analysis to enhance accuracy and provide a more detailed characterization of detected flaws.
2. Developing temperature-compensated calibration standards specifically designed for PACM inspections would improve consistency and reliability across different operational temperatures.
3. Conducting field trials in operational petrochemical plants to validate laboratory findings under actual industrial conditions and demonstrate the practical benefits of PACM in reducing downtime and maintenance costs.

By addressing these recommendations, future research can further establish PACM as a cornerstone technology for corrosion detection in high-temperature industrial applications, ultimately contributing to enhanced operational safety, extended equipment life, and reduced maintenance costs in the petrochemical and related industries.

ACKNOWLEDGEMENT

The authors are grateful for the financial support given by the Ministry of Higher Education Malaysia (MOHE) under the Higher Institution Centre of Excellence (HICOE2.0/5210004) at the Institute of Tropical Forestry and Forest Products. 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 Products (INTROP), Universiti Putra Malaysia (HICOE) for their close collaboration in this study.

REFERENCES

Aiello, G., Ben, J., Carpitella, S., Certa, A., Enea, M., & La Cascia, M. (2020). A decision support system to assure high-performance maintenance service. *Journal of Quality in Maintenance Engineering, 27*(4), 651-670. https://doi.org/10.1108/JQME-11-2019-0107

Al-Marri, A. N., Nechi, S., Ben-Ayed, O., & Charfeddine, L. (2020). Analysis of the performance of TAM in oil and gas industry: Factors and solutions for improvement. *Energy Reports*, *6*, 2276–2287. https://doi.org/10.1016/j.egyr.2020.08.012

Al-Turki, U., Duffuaa, S., & Bendaya, M. (2019). Trends in turnaround maintenance planning: Literature review. *Journal of Quality in Maintenance Engineering, 25*(2), 253-271. https://doi.org/10.1108/JQME-10-2017-0074

Annis, C. (2023). *Statistical Engineering*. (2023). https://statistical-engineering.com/

ASTM E3023-21. (2021). *Standard Practice for Probability of Detection Analysis for a Versus a Data*. ASTM International. https://store.astm.org/e3023-21.html

ASTM E2862-23. (2023). *Standard Practice for Probability of Detection Analysis for Hit/Miss Data*. ASTM International. https://store.astm.org/e2862-23.html

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

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*, Article 102300. https://doi.org/10.1016/j.ndteint.2020.102300

Cherry, M., & Knott, C. (2022). What is probability of detection? *Materials Evaluation*, *80*(12), 24–28. https://doi.org/doi.org/10.32548/2022.me-04324

Department of Defense Handbook. (2009). *Nondestructive evaluation system reliability assessment*. Department of Defense Handbook. https://statistical-engineering.com/wp-content/uploads/2017/10/MIL-HDBK-1823A2009.pdf

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

Elwerfalli, A., Khan, M. K., & Munive-Hernandez, J. E. (2019). Developing turnaround maintenance (TAM) model to optimize TAM performance based on the critical static equipment (CSE) of GAS plants. *International Journal of Industrial Engineering and Operations Management*, *01*(01), 12–31. https://doi.org/10.46254/j.ieom.20190102

Goursolle, T., Fauret, T., & Juliac, E. (2016). Effect of data amount on probability of detection estimation: Application to Eddy current testing. *e-Journal of Nondestructive Testing, 21*(7), 1-8. http://creativecommons.org/licenses/by/3.0/

Hlophe, S. C., & Visser, J. K. (2018). Risk management during outage projects at power plants. *South African Journal of Industrial Engineering*, *29*(3), 82–91. https://doi.org/10.7166/29-3-2051

Jory, C. (2019). Tips for internal corrosion using the echo to echo technique with compression. *The NDT Technician*, *18*(3), 8-11.

Keprate, A., & Ratnayake, R. (2015). Probability of detection as a metric for quantifying NDE capability: The state of the art. *Journal of Pipeline Engineering*, *14*(3), 2015–2017.

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., Ciarallo, F., & V.Grandhi, R. (2019). Developments in probability of detection modeling and simulation studies. *Materials Evaluation*, *73*(1), 55–61.

Knott, C. E., & Kabban, C. S. (2022a). Confidence interval comparisons for probability of detection on hit/miss data. *Materials Evaluation*, *80*(12), 50–65. https://doi.org/doi.org/10.32548/2022.me-04273

Knott, C. E., & Kabban, C. S. (2022b). Modern design and analysis for hit/miss probability of detection studies using profile likelihood ratio confidence intervals. *Materials Evaluation*, *80*(12), 32–49. https://doi.org/doi.org/10.32548/2022.me-04272

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.

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. https://source.asnt.org/nondestructive-examination-(nde)-of-used-fuel-containers-for-increased-probability-of-detection/

Rentala, V. K., & Kanzler, D. (2022, October 24-27). Theoretical POD Assessment of an NDE 4.0 Application under the Context of Aero-Engine Lifing. In *Proceedings of the 2nd International Conference on NDE* (Vol. 4). Berlin, Germany. https://conference.nde40.com/Portals/NDE40_2021/bb/We.1.B.1.pdf

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

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 *7th European-American Workshop on Reliability of NDE* (Vol. 22, No. 12). 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

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

Tai, J. L., Sultan, M. T. H., Tarasiuk, W., Napiórkowski, J., Łukaszewicz, A., & Shahar, F. S. (2023). Ultrasonic velocity and attenuation of low-carbon steel at high temperatures. *Materials*, *16*(14), Article 5123. https://doi.org/10.3390/ma16145123

Tai, J. L., Sultan, M. T. H., Shahar, F. S., Łukaszewicz, A., Oksiuta, Z., & Grzejda, R. (2024). Ultrasound corrosion mapping on hot stainless steel surfaces. *Metals*, *14*(12), 1–16. https://doi.org/10.3390/met14121425

Tai, J. L., Sultan, M. T. H., Shahar, F. S., Yidris, N., Basri, A. A., & Shah, A. U. M. (2024). Exploring probability of detection (POD) analysis in nondestructive testing: A comprehensive review and potential applications in phased array ultrasonic corrosion mapping. *Pertanika Journal of Science and Technology*, *32*(5), 2165–2191. https://doi.org/10.47836/pjst.32.5.14

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*, Article 89. https://doi.org/10.1007/s10921-019-0628-z

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., 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

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