Improvement of ILI Sizing Accuracy for Problematic Corrosion Profiles

Pablo Cazenave, Katina Jimenez and Ravi Krishnamurthy
Blade Energy Partners Ltd.

PPIM 2025
PIPELINE PIGGING & INTEGRITY MANAGEMENT
JANUARY 27-31 • HOUSTON

Organized by
CLARION
TECHNICAL CONFERENCES

https://doi.org/10.52202/078572-0074

Abstract

In March 2019, the U.S. Pipeline and Hazardous Materials Safety Administration (PHMSA) issued a Research Announcement to improve the detection and sizing accuracy of In-Line Inspection (ILI) systems for critical pipeline anomalies, aiming to improve safety and reduce unnecessary excavations. In response, Pipeline Research Council International (PRCI) launched Project NDE-4-19 in December 2019, a multi-phase initiative in collaboration with PHMSA and Blade Energy Partners. The project brought together Pipeline Operators, ILI Technology Providers (TPs), and Subject Matter Experts to assess and improve ILI system capabilities for challenging corrosion features. A data-driven approach was used, in which recent Root-Cause-Analysis (RCA) reports of corrosion-related pipeline failures were analyzed, and a pipe test string was designed with complex corrosion profiles resembling those associated with the failures. The manufactured corrosion features were documented using advanced non-destructive evaluation (NDE) techniques.

Three ILI TPs tested their tools on the designed pipe test string in a series of blind pull-through tests. Following initial evaluations, the TPs received feedback on detection and sizing gaps and were provided with detailed 3D corrosion profiles to guide their improvements. A second round of tests measured the enhancements in detection and sizing accuracy, identifying improvements and remaining challenges.

This paper presents the findings from this 3.5-year-long project, highlighting advancements in ILI technology and ongoing gaps in detecting and sizing problematic corrosion profiles.

Keywords: Pipeline Integrity, ILI Validation, ILI Improvement, ILI detection capability, ILI sizing accuracy, corrosion integrity evaluation.

Introduction

In September 2018, a Pipeline Research and Development Forum in Baltimore, MD, gathered industry experts and government officials to tackle key challenges in pipeline safety. One of the main priorities identified was the need to improve how pipeline anomalies are detected, sized, and characterized.

By March 2019, the Pipeline and Hazardous Materials Safety Administration (PHMSA) issued a call for action to advance In-Line Inspection (ILI) technology. The aim was to enhance detection and sizing accuracy, making detecting pipeline anomalies easier, reducing unnecessary excavations, and concentrating efforts on the most critical safety threats.

This initiative is designed to push ILI systems toward a goal of 100% accuracy in identifying critical conditions. Achieving this will improve pipeline safety and ensure more efficient use of resources and less disruption during maintenance activities.

Background

Pull-through test facilities for in-line inspection (ILI) systems are rare and typically owned by ILI

https://doi.org/10.52202/078572-0074

technology providers (ILIs). These facilities are crucial for testing tool performance, calibrating inspection tools, and refining signal analysis algorithms. However, independent pull-through facilities are not widely accessible to operators.

PRCI owns and operates a large-scale pull-through facility at its Technology Development Center (TDC), available to ILI technology providers and pipeline operators. The TDC helps advance ILI technologies by evaluating their ability to identify pipeline features, anomalies, and defects.

In 2015, PRCI and PHMSA's NDE-4F project[1,2] focused on enhancing ILI technology evaluation and identifying performance gaps. The project tested commercially available Magnetic Flux Leakage (MFL) technologies and found that large features (over 10A in the POF[3] diagram) often showed the highest depth results outside ILI specifications—a well-known issue in the MFL inspection industry. It also highlighted challenges in accurately measuring total feature lengths, with similar variability observed in corrosion[1], dents[4,5], and other defects[6,7]. Additionally, NDE-4F identified that pinholes in areas of corrosion and widespread wall loss are the most concerning defects. It recommended using advanced measuring techniques for reference data[4,8] and comparing burst pressure estimates via effective length[9,10] rather than total length[11]. However, the industry still lacks a complete data-driven understanding of the defect profiles responsible for recent corrosion failures.

Objectives

This PRCI/PHMSA-sponsored study brings together pipeline operators, technology providers (ILI vendors), subject matter experts, and consultants to enhance the detection and sizing capabilities of current ILI technologies. Building on past PRCI/PHMSA research[2], this collaborative, multi-phase project addresses gaps using industry-relevant, fact-based data, and quantitative analysis.

Focusing on corrosion—one of the largest threats to pipeline safety—the study highlights that corrosion and mechanical damage are the leading causes of leaks and ruptures in North American transmission pipelines. With 80% of ILI runs and 70% of excavations linked to corrosion[12], the study aims to improve ILI sizing accuracy for complex corrosion anomalies, particularly small pits, pinholes, deep depths, and wide-area wall loss, and to provide technology developers with high-quality, high-resolution, three-dimensional reference data to refine sensing capabilities and sizing algorithms using real-life corrosion data.

Why Do We Still Have Failures Due to Corrosion?

ILI technology for corrosion features has been extensively developed over decades, with billions of dollars invested. However, failures—primarily leaks and ruptures—still occur despite using advanced ILI tools. To understand the corrosion profiles causing these failures, we gathered fact-based data from 33 pipeline operators (17 liquid and 16 gas) through detailed root cause analysis (RCA) reports and associated ILI data. We obtained over 25 RCA reports, with 19 providing detailed information on the corrosion features involved in failures.

https://doi.org/10.52202/078572-0074

These corrosion defects were classified based on their shape, distribution, and association with other features like welds. Examples of these profiles are shown in Figures 1, 2, and 3.

Figure 1: Leaking corrosion feature

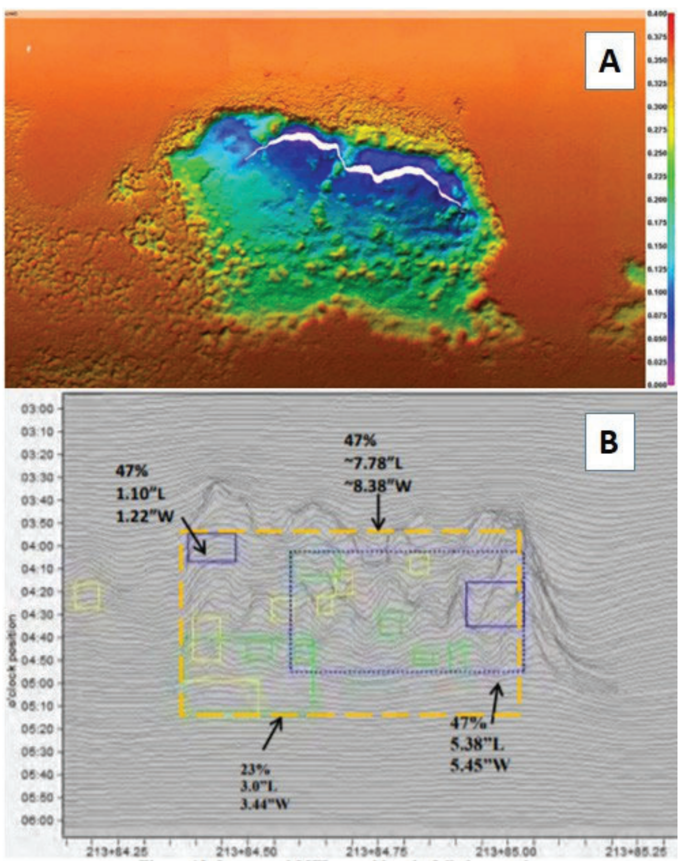

Figure 2: (a) Ruptured corrosion feature (b) MFL signal data of the same feature before the failure occurred

https://doi.org/10.52202/078572-0074

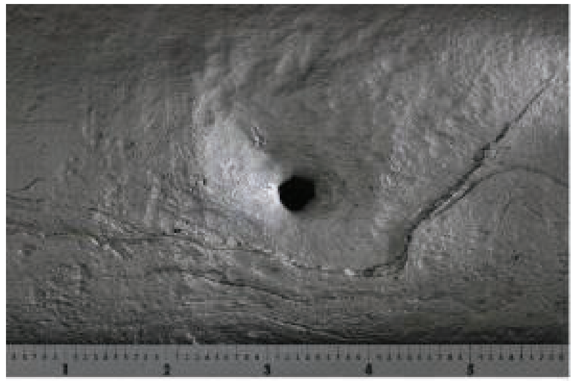

Figure 3: Corrosion feature that leaked

The study identified several problematic corrosion features that caused failures but were either undetected or undersized by recent ILIs, namely flat-bottom corrosion, corrosion across girth welds, general corrosion, axial slit leaks (long axial grooves), funnel-shaped corrosion, isolated pits, wider circumferential corrosion, longer axial corrosion, circular corrosion, corrosion near bends.

Further discussions with ILI Technology Providers and Subject Matter Experts highlighted other challenging corrosion profiles such as short, deep pinhole corrosion, extensive general corrosion or wall thinning over large areas, axial corrosion in specific areas (e.g., bottom of the pipe, gas-liquid interface), corrosion with coincidental internal and external occurrences, defects at angles to the pipe axis, and smooth, gradual thinning over large areas.

These corrosion profiles are the focus of the current study.

Finding pipe with the target anomalies

One of the most challenging and time-consuming tasks in ILI testing is finding pipe samples with natural corrosion anomalies grouped in a single diameter suitable for testing. To design this experiment, Blade surveyed all available pipe samples with potential target anomalies, focusing on selecting corrosion profiles for ILI testing. This process included scouting, conducting preliminary NDE to identify suitable corrosion features, building a pipe string with these anomalies, gathering high-quality reference data, and accounting for statistical and facility constraints.

The survey targeted pipes with corrosion defects at the TDC, collaborating with other PRCI projects like NDE 4-18[13], and reviewing numerous 20-inch samples. The surveyed pipes contained various corrosion features, including those with repaired areas under steel sleeves. The sleeves were removed to expose and catalog the corrosion features, following a protocol developed in the early stages of PRCI Project MD-1-13[14] to preserve the anomalies.

The sleeve removal process exposed deep and complex corrosion features, providing natural corrosion profiles for the project (see example of complex corrosion profiles in Figure 4).

https://doi.org/10.52202/078572-0074

Figure 4: Example of corrosion features recovered

However, the pipe samples lacked certain specific target profiles. To fill these gaps, additional joints were created with features like corrosion on girth welds, smooth thinning, and coincidental external/internal corrosion. These profiles were created manually using grinding, buffing, and sanding methods since anodic dissolution or acid etching could not achieve the desired metal loss profiles. The grinding process did create a cold-worked layer, which was then removed by etching or sanding, as confirmed through microscopic examination (see Figure 5).

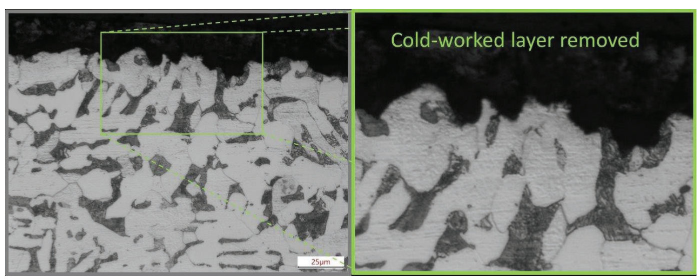

Figure 5: Example of superficial cold-worked layer removed

Finally, all joints with defects were 3D laser scanned, documented, and coated with corrosion-resistant enamel to prevent atmospheric degradation while stored outdoors. The pipe string included a mix of seam-welded (ERW and DSAW) and seamless pipes. Approximately 75% of the metal loss features used in this string were naturally occurring corrosion features recovered from operational pipelines.

ILI Testing and Improvement Approach

The ILI testing approach followed a four-step process:

1. Initial blind ILI test (T1): This identified the sizing strengths and gaps of each ILI system.
2. Corrosion profile feedback: A small number of features were shared with ILI technology providers (ILIs) for analysis.
3. System improvements: The ILI TPs identified and implemented necessary improvements.
4. Retesting (T2): The improved ILI system was retested to evaluate performance.

https://doi.org/10.52202/078572-0074

The ILI technologies evaluated included an Ultrasonic (UT, 90° Wall Thickness measurement) + Caliper (identified here as ILI TP 1), an Axial Magnetic Flux Leakage (MFL, triaxial Hall sensors) + Caliper (ILI TP 2), and an Axial Magnetic Flux Leakage (MFL, single-axis Hall sensors) + Caliper (ILI TP 3). For simplicity, these were referred to as UT for ILI TP1 and MFL for ILI TP2 and TP3.

The Reference NDE data were obtained by measuring the metal loss 3D profiles. All testing joints were 360° laser-scanned in the same orientation as the expected ILI runs. Peak depth, total length, and total width of significant features were verified using manual measurements (pit gauge micrometer, caliper, measuring tape, and ultrasonic thickness gauge).

ILI Test Runs in Water (Push-Through)

A U-shaped water push-through string was built for testing Ultrasonic In-Line Inspection systems (UT ILI), which require a liquid medium as a coupling agent between the sensors and pipe wall. The string included joints with mechanical damage (MD) features for another project, and the natural corrosion and synthetic metal loss (ML) features suitable for this study.

The water test string's total length, excluding launching and receiving joints, was approximately 973 ft (297 m). Figure 6 illustrates the layout of the water test string at the TDC.

Figure 6: Water String

ILI Test Runs in Air (Pull-Through)

A pull-through test string was designed for ILI tools that do not require a liquid couplant; Figure 7 shows the layout of the pull-through string at the TDC.

https://doi.org/10.52202/078572-0074

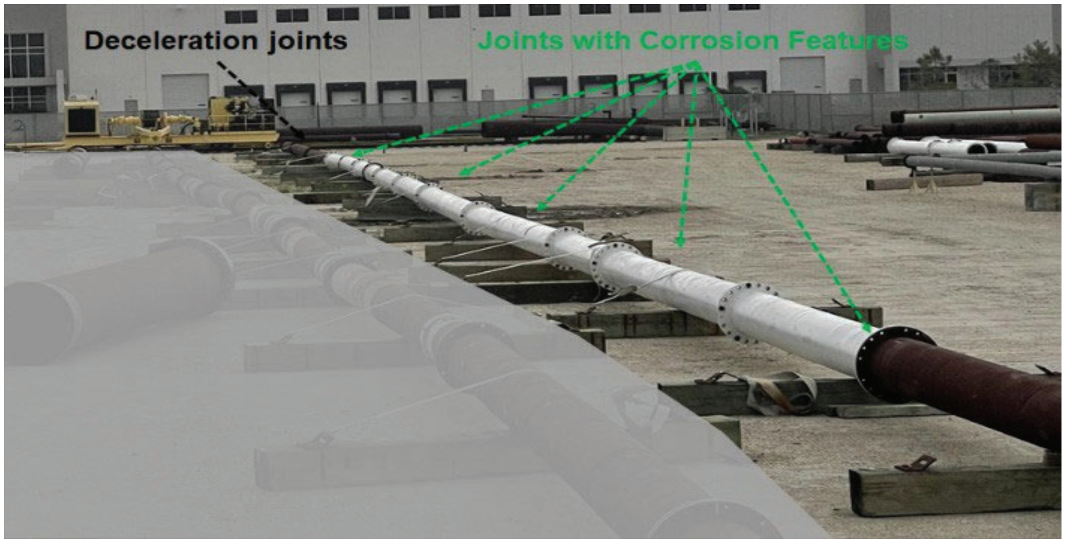

Figure 7: Pull-through string

Feature matching and Performance Evaluation

The features reported by the ILI were manually aligned and matched with their Reference NDE counterparts for analysis, following industry best practices. Experienced personnel, familiar with both the capabilities and the limitations of the ILI systems and the NDE methods used for the reference features, carried out the matching process. This understanding of both ILI and NDE specifications and limitations is essential for accurate analysis and effective statistical evaluation.

The results were then analyzed to identify areas where the ILI system could be improved. The evaluation compared calculated burst pressure using effective area methodologies (e.g., R-Streng Effective Area) between the ILI and Reference data (rupture condition). Similarly, peak depth evaluations for features that failed by leak were compared (leak condition).

Three principal measures are typically used to evaluate ILI system performance: Probability of Detection (POD), Probability of Identification (POI), and Sizing Accuracy (error). POD measures the ability to detect an anomaly, POI measures the ability to identify its type, and sizing accuracy evaluates the system's ability to accurately report the anomaly's dimensions (depth, length, and width).

Although methods like binomial confidence intervals[15,16], Clopper-Pearson[16,17] and Agresti-Coull[18] are often used to estimate POD, POI, and sizing accuracy from limited sample sizes, the entire population of features in the test string was known. Therefore, probabilities were calculated as the ratio of correct calls to total observations.

Sizing accuracy is often reported as tolerance and confidence level, such as ±10% at 80% confidence, and is visualized using unity plots to show the sizing accuracy and tolerance.
After analyzing the performance challenges for each ILI system, a detailed data package was prepared

 https://doi.org/10.52202/078572-0074

for each ILI TP. This package included a summary of findings, overall results, unity plots, 3D profile screenshots of problematic features, and their relative positions in the unity plots. Between 26 and 40 features (out of a population of approximately 1000) were provided to each ILI TP, while the rest remained confidential for Test Trial 2. The feedback data also included 3D profiles (2 x 2 mm grid) in spreadsheet format.

The three participating ILI TPs reviewed the feedback. They identified improvements in sensor interrogation, signal processing, and sizing algorithms to enhance the detection, discrimination, and sizing of these challenging corrosion features. They concluded that their existing hardware configurations (tools, sensors, sensor carriers, onboard electronics) were sufficient for corrosion detection and characterization and that modifying signal processing and software analysis would improve defect detection and sizing.

Retesting the Improved ILI System

The project's final phase focused on retesting the ILI systems to evaluate improvements and identify remaining detection and sizing gaps. During this phase, the improved ILI system was retested (ILI Test #2, or T2) by reanalyzing the signal data collected during the initial blind test (T1). The hardware system remained unchanged.

For each problematic corrosion profile, improvements were *quantified* in terms of:

- *Actionable conditions captured by the ILI*: This refers to leak and rupture conditions identified by the ILI. A condition is considered captured when the ILI reports a rupture or leak based on the metal loss feature sizing in the Reference data. A rupture is considered captured when both the Reference and ILI Estimated Repair Factor (ERF) are greater than or equal to 1. A leak is considered captured when the maximum depth (dmax) in both the Reference and ILI data is at least 80% of the pipe wall thickness.
- *False positives*: This refers to actionable conditions unnecessarily reported by the ILI.
- *Precision and accuracy*: Accuracy measures how close measurements are to the actual value, while precision measures how consistent the measurements are.

Figures 9, 10 and illustrate these comparisons for all corrosion and metal loss features in the string for ILI TPs 1 and 3.

https://doi.org/10.52202/078572-0074

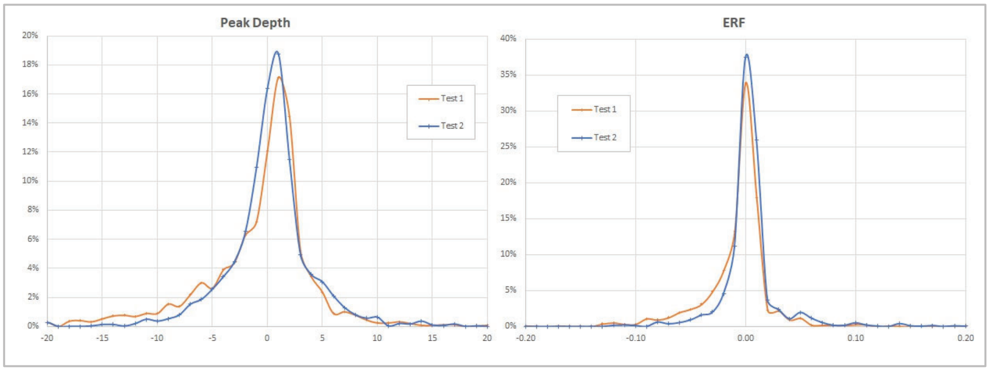

Figure 8: Relative frequency distributions for d_{max} and ERF differences in T1 and T2 – ILI TP 1

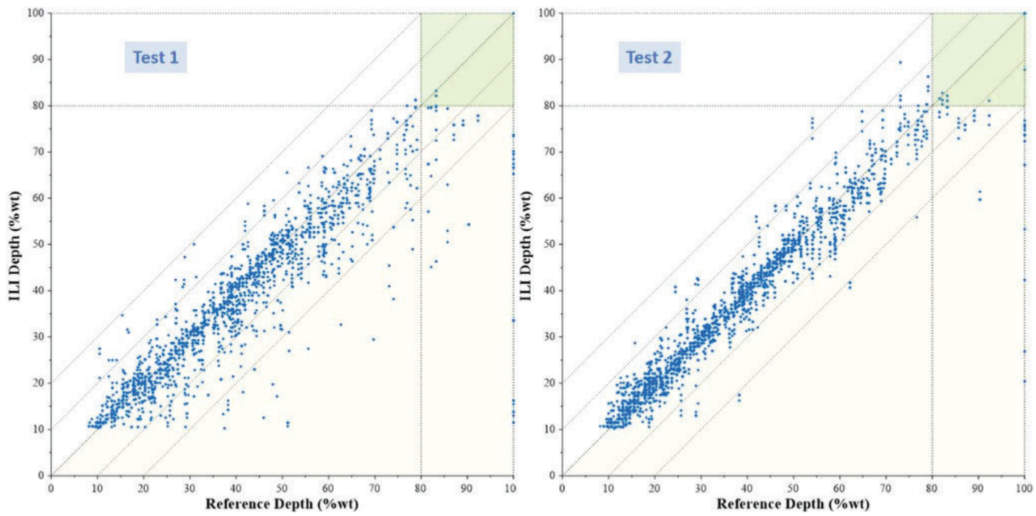

Figure 9: Unity plots for d_{max} as %t in T1 and T2 – ILI TP1

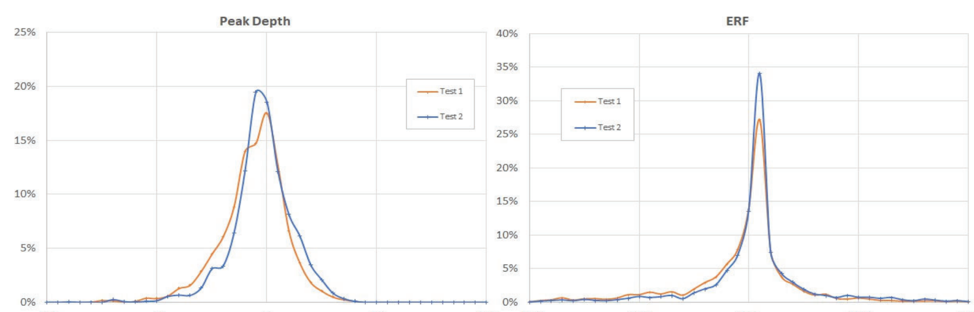

Figure 10: Relative frequency distributions for d_{max} and ERF differences in T1 and T2 – ILI TP3

https://doi.org/10.52202/078572-0074

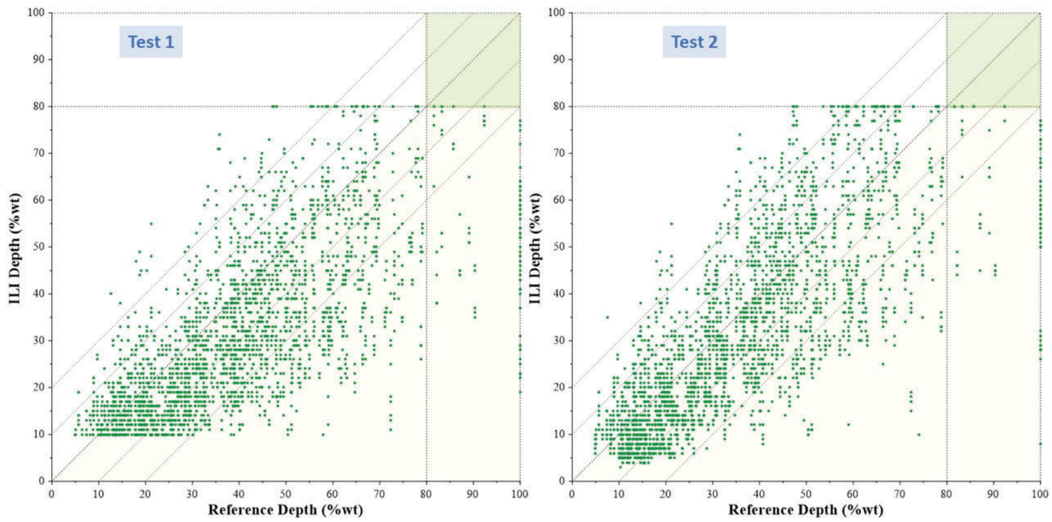

Figure 11: Unity plots for d_{max} as %t in T1 and T2 – ILI TP3

Results and discussion

It should be noted that most of the metal loss (ML) features in this test string have complex and challenging profiles, many of which fall outside the commercial specifications of the evaluated ILI systems. Several ML features have dimensions that make them undetectable or challenging to size. Therefore, the absence of 100% Probability of Detection (POD) and a high-depth scatter was expected.

ILI Test #1 results showed that MFL systems captured more leak conditions, especially for short and narrow pits, while UT systems captured more rupture conditions than MFL systems. However, not all problematic profiles proved difficult for the ILI technologies. After Test #1, the most consistently challenging corrosion profiles for all evaluated ILI technologies included:

- Axially oriented narrow features
- Deep pitting within general corrosion
- Features with flat bottom profiles
- Features with smooth, gradual corrosion depth changes
- Features crossing girth welds
- Pinholes and short, narrow pits
- Funnel-shaped pitting

Additionally, the categories most prone to false positives included:

- Clustering issues in areas of general corrosion with pitting, where ILI grouped areas differently from the Reference NDE method
- Circumferentially oriented features
- Smooth and gradual wall thinning

https://doi.org/10.52202/078572-0074

After ILI Test #2, two of the three improved ILI systems (UT and MFL TP3) significantly improved sizing and detection. However, MFL TP2 did not achieve comparable detection and sizing improvements but substantially reduced the number of false positives.

The most notable improvements for the UT ILI system were observed in Flat bottom profiles, Narrow Axial Extended Corrosion (NAEC), General corrosion with pitting, and Corrosion with a wider circumferential extent. Conversely, Funnel-shaped corrosion (pitting/pinhole) showed the least improvement.

For MFL TP3, significant improvements were seen in Smooth and gradual wall thinning, General corrosion with pitting, Axially oriented NAEC, Funnel-shaped corrosion, and Flat bottom profiles, while the categories showing less improvement included Short, deep corrosion (pinhole), Coincidental internal and external corrosion, and Pitting corrosion with longer axial extent.

It is also important to note that ILI systems 1 (UT) and 3 (MFL) detected all features with the most significant rupture conditions (calculated failure pressure < MAOP), while system 2 (MFL) captured most of them.

While leak/burst condition identification improvements were achieved, there was also an increase in false positives for some cases. The ideal goal is to capture more actionable conditions while minimizing unnecessary conditions, which was achieved for specific ML feature categories.

Despite these improvements, the ILI systems still missed some leaking conditions, as these remain the most challenging for both ILI and NDE systems. Future projects testing new or improved ILI technologies could address these challenges and achieve better results.

At the conclusion of the project, the ILI technology providers received a comprehensive dataset, including all matched ML features, 3D defect profiles in universal Excel format, and full-joint 3D laser scan files. This extensive dataset (approx. 1,000 ML features) provides them with valuable insights to further enhance the detection, sizing, and characterization capabilities of their ILI systems, which is crucial for improving pipeline safety. PRCI has preserved the documented strings for future testing, and the strings and approach used in this project can be expanded to other ILI providers and ILI systems.

Conclusions

The most critical and effort-consuming stage in any ILI testing exercise is acquiring sufficient pipe samples with natural (non-synthetic) features that are feasible for testing. In this project, pipeline operators contributed pipes with significant corrosion that were used for the testing. Such samples' continued collection and availability are essential for ongoing improvements in ILI technologies.

The results revealed that MFL systems were more effective at capturing leak conditions, especially for

https://doi.org/10.52202/078572-0074

short and narrow pits, while UT systems were better at capturing rupture conditions. Statistical analysis of system performance showed that ILI Systems 1 (UT) and 3 (MFL) successfully identified all features with the most significant rupture conditions (FP < MAOP), whereas ILI System 2 (MFL) captured most of the rupture conditions.

The project demonstrated that ILI system performance can be improved when high-quality and detailed reference data is provided as feedback to the ILI Technology Providers (TPs). This feedback loop is crucial for refining ILI systems' detection, sizing, and characterization capabilities.

The fully documented 20-inch corrosion string created during the project is available for further assessment and improvement of ILI systems, including those that operate beyond their current specifications. Emerging systems, such as High-Sensor-Density (HD MFL), Ultra-High-Resolution (UHR) Axial and Circumferential MFL systems[19,20], and ILI technologies based on acoustic resonance[21], may be viable candidates for addressing the limitations identified in this study.

Due to space limitations, this paper only presents a limited number of comparisons, but the detailed quantitative results (including tabular data and graphs) for each corrosion profile category are available in the final PRCI/PHMSA report, which can provide deeper insights into each ILI system's performance across the various feature types.

Acknowledgments

The authors would like to express their gratitude to the PRCI Project Team Leaders for their significant contribution to the preparation of this study. Special thanks are extended to the participating ILI technology providers, the PRCI Team, and the pipeline operators for their continued support and assistance throughout the execution of this project.

The authors would also like to acknowledge the co-sponsorship and support provided by PHMSA. Their involvement, including reviewing interim reports and providing invaluable questions and comments, has played a crucial role in shaping the outcomes of this project over the course of more than three years of execution.

Disclaimer

This research was partly funded under the Department of Transportation, Pipeline and Hazardous Materials Safety Administration’s Pipeline Safety Research and Development Program. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Pipeline and Hazardous Materials Safety Administration or the U.S. Government.

https://doi.org/10.52202/078572-0074

References

1. P. Cazenave and M. Gao, PHMSA/PRCI Project, "Development of an industry test facility and qualification process for ILI technology and Evaluation Enhancements - ILI Trials and Performance Evaluation Phase". PRCI Project NDE-4F, PR-328-153721 / PHMSA Contract DTPH56-13-T-000010. February 21st, 2017.
2. P. Cazenave, M. Gao, H. Deeb and S. Black. "Development of an Industry Test Facility and Qualification Process for ILI Technology Evaluation and Enhancements - Performance Evaluation Phase". Proceedings of the 12th International Pipeline Conference, Paper IPC2018-78635. September 24-28, 2018, Calgary, Alberta, Canada.
3. Pipeline Operators Forum. "Specifications and requirements for in-line inspection of pipelines". Standard Practice POF 100. November 2021.
4. Blade Energy Partners. "Field Testing and Verification of Existing Tool Capabilities for Mechanical Damage Detection and Characterization". PRCI Project (MD-1-2), PR-328-073511, June 2013.
5. P. Cazenave, K. Jimenez and M. Gao. "Evaluation of ILI Capabilities on Mechanical Damage Features". PRCI project MD-1-13, PR-328-173865, May 2020.
6. P. Cazenave, K. Jimenez and M. Gao., Blade Energy Partners. "Performance Capabilities Evaluation of ILI for ERW Pipe, Long Seam Features via water-loop testing at PRCI's Technology Development Center". PRCI project IM-3-1, PR-328-173866, September 2019.
7. M. Gao et al. "Continuous Evaluation of EMAT Tool Performance and Reliability by Monitoring Industry Experience". PR-328-083501-R01, 2015.
8. P. Cazenave, K. Tiñacos-Jimenez, M. Gao et al. "Evaluation of new in-ditch methods for measurement and assessment of external corrosion". Proceedings of the 10th International Pipeline Conference, Paper IPC2014-33578. September 29 - October 3, 2014, Calgary, Alberta, Canada.
9. American Society of Mechanical Engineers, ASME. "ASME B31G-2009 - Manual for determining the remaining strength of corroded pipelines". Section 2.3, Level 2 Evaluation, Effective Area Method or RStreng, iterative method. 2009.
10. Det Norske Veritas. "DNV Recommended Practice RP-F101 – Corroded Pipelines". Iterative method. January 2015.
11. American Society of Mechanical Engineers, ASME. "ASME B31G-2009 - Manual for determining the remaining strength of corroded pipelines". Sections 2.1 (Level 0 evaluation) and 2.2 (Level 1 evaluation). 2009.
12. S. Kariyawasam et al., "Plausible Profiles (Psqr) Model for Corrosion Assessment." Technical Report Prepared by TC Energy for the Corrosion Technical Committee of the PRCI. July 22nd, 2019.
13. PRCI Project NDE-4-18. "Performance of ILI Technologies for Dents with Interacting Features". Ongoing research project.
14. P. Cazenave, K. Jimenez and M. Gao. "Evaluation of ILI Capabilities on Mechanical Damage Features". PRCI Project MD-1-13, PR-328-173865, May 2020.
15. R. McCann, R. McNealy, and M. Gao. "In-Line Inspection Performance Verification". NACE Corrosion 2007, Paper 07132. 2007.
16. M. Gao, R. Krishnamurthy: "A Review: Statistical Methods for In-Line Inspection Performance Evaluation", Rio Pipeline Conference & Exposition 2011, Paper No: IBP2093-11, September 20-22, 2011, Rio de Janeiro, Brazil.
17. R. McCann, R. McNealy, M. Gao. "In-Line Inspection Performance II – Validation sampling". NACE Corrosion 2008, Paper 08151. 2008.
18. PRCI Project IM-6-1, "API-1163 Section 8 - Level 1, 2 & 3 Research".
19. R. Desaulniers, "Applying Ultra-High-Resolution MFL to Achieve a Better Integrity Assessment of Pits-in-Pits." Proceedings of the 35th Pipeline Pigging and Integrity Management Conference. Houston, TX, USA, February 6-10, 2023.
20. R. Thompson, A. Corbett, G. Solano, and R. Kania. "Integrity Management of Off-axis SCC Through and ILI System and Engineering Assessment." Proceedings of the 36th Pipeline Pigging and Integrity Management Conference. Houston, TX, USA, February 12-16, 2024.
21. W. Vos, J. Ove Toskedal and G. Dye. "Review of 10 Years of Acoustic Resonance Applications for Pipeline In-Line Inspection". Proceedings of the 36th Pipeline Pigging and Integrity Management Conference. Houston, TX, USA, February 12-16, 2024.

 https://doi.org/10.52202/078572-0074

https://doi.org/10.52202/078572-0074