

CAAP Final Report

Date of Report: December 4, 2023

Prepared for: *U.S. DOT Pipeline and Hazardous Materials Safety Administration*

Contract Number: *693JK31950004CAAP*

Project Title: Multi-modal NDE Assisted Probabilistic Pipeline Performance Evaluation under
Interactive Anomalies

Prepared by: *Qindan Huang, Qixin Zhou, and YimingDeng*

Contact Information:

Qindan Huang

Associate Professor

Civil, Construction & Environmental Engineering Department

Marquette University

1637 W. Wisconsin Ave., Milwaukee, Wisconsin 53233

Phone: 414-288-6670

Multi-modal NDE Assisted Probabilistic Pipeline Performance Evaluation under Interactive Anomalies

Table of Contents

1	Executive Summary.....	9
1.1	Summary of Accomplishments.....	10
2	Introduction/Background and Objectives.....	12
2.1	Objectives	13
2.2	Justification of Scope Adjustment	14
3	Objective 1: Lab Testing of Generating Realistic Corrosion Defect.....	15
3.1	Experimental Program	15
3.2	Results and Discussion.....	15
4	Objective 2: NDE Framework.....	23
4.1	Eddy Current Array Methods for Interactive Corrosion Detection and Characterization	23
4.2	Ultrasonic NDE methods for interactive corrosion detection and characterization.....	30
4.3	Dataset Preparation for Machine Learning based Data Analysis.....	60
4.4	Ultrasonic Imaging Methods.....	70
5	Objective 3: Probabilistic models of failure pressure prediction	76
5.1	Models for Pipeline with Single Corrosion Defect.....	76
5.2	Probabilistic Models of Defect Interaction Identification and Failure Pressure for Pipelines with Interacted Corrosion Defects	101
5.3	Models for Pipeline with Single Crack-like Defect	122
6	Objective 4: Uncertainty Impact on Pipeline Reliability.....	140
6.1	Background.....	140
6.2	Case Study I.....	140
6.3	Case Study II.....	145
6.4	Case Study III	147
7	Conclusions	152
7.1	Net Safety Impact	156
7.2	Future Work	156
	References.....	158
	Appendix A: Exposure Testing Results of Task 1.....	164
	Appendix B: Existing models of burst failure prediction for pipeline with single corrosion defect	
	169	

Figure 1. Quantitative risk management for pipeline.	12
Figure 2. Infinite microscopy images of the flat testing samples during B117 exposure testing.	18
Figure 3. The sample mean and sample standard deviation of the average depth and the maximum depth of the scanning area of all the flat samples during B117 exposure testing.	19
Figure 4. The sample mean and sample standard deviation of the average depth and the maximum depth of the scanning area of all the flat samples during G85 exposure testing.	19
Figure 5. The depth of the pre-damaged testing sample continuously under B117 exposure.	20
Figure 6. The depth of the pre-damaged testing samples under B117 exposure. Orange: initial depth; green: the depth at the time of testing.	20
Figure 7. The change of the depth of the pre-damaged testing samples under B117 exposure.	21
Figure 8. The depth of the pre-damaged testing sample continuously under G85 exposure.	21
Figure 9. The depth of the pre-damaged testing samples under G85 exposure. Orange: initial depth; green: the depth at the time of testing.	22
Figure 10. The change of the depth of the pre-damaged testing samples under G85 exposure. ..	22
Figure 11. Corroded samples and their respective labels.	26
Figure 12. 2D post-processed data from real voltage component on corroded samples using the left column of coils.	26
Figure 13. 2D post-processed data from real voltage component on corroded samples using the right column of coils.	27
Figure 14. Eddy Current Array (ECA) NDE gantry setup.	28
Figure 15. 1D post-processed data from real voltage component on corroded samples using coil 16.	29
Figure 16. Dispersion curve for Steel AISI 4340	32
Figure 17. Sample geometry for 2-D axisymmetric model to simulate axial waves in pipes.	33
Figure 18. Time domain representation of the burst signal	34
Figure 19. Frequency domain representation of the burst signal.	34
Figure 20. Zoomed simulation model showing the excitation by the application of the burst type signal on transducer boundary.	35
Figure 21. Velocity profile at (0,0) of the sample.	36
Figure 22. Velocity profile at (0,0) and (0,0.5) of the sample	37
Figure 23. Von Misses Stress at time $t=0.002$ seconds	37
Figure 24. Cross section of the pipeline with a small pit.	39
Figure 25. (a) A-scan at (0,0) and (b) Von Misses stress at time $t = 0.002s$	39
Figure 26. Comparison of A-scans with baseline	40
Figure 27. Defect Signature for pit with 1 mm depth	41
Figure 28. A-scans and defect signature for pit with 3 mm depth.	41
Figure 29. A-scans and defect signature for pit with 5 mm depth.	42
Figure 30. Defect signatures for pits with 3mm and 5 mm depths	42
Figure 31. 2 pits modelled adjacent to each other on the surface	43
Figure 32. A-scans and defect Signature for two pits with 3 mm depth.	43
Figure 33. Defect signatures for single and double pit models.	44
Figure 34. (a) A-scan at (0,0) and (b) von misses stress at time $t = 0.002s$	44
Figure 35. Comparison of A-scans with baseline	45

Figure 36. Defect Signature for pit with 1 mm depth	45
Figure 37. Defect signatures for pits with 3mm and 5 mm depths	46
Figure 38. Defect signatures for single and double pit models.....	46
Figure 39. (a) Corrosion type defects located circumferentially around the pipe, (b) zoomed image showing dimensions of a single pit, (c) colony of three pits, and (d) interacting defects consisting of a colony of three pits with 2 mm deep crack in the central pit.....	48
Figure 40. Circumferential guided wave around the pipe seen from the von Mises stress profile across the cross section.	49
Figure 41. Axial displacement response at the piezoelectric wafer transducer location with healthy, one pit, two pit colony and three pit colony conditions.	49
Figure 42. Axial displacement response at the piezoelectric wafer transducer location with three pit colony, interactive 3 pit colony with 1 mm crack, interactive 3 pit colony with 2 mm crack and interactive 3 pit colony with 3 mm crack conditions.	50
Figure 43. Split of the 150 samples collected for Classification	52
Figure 44. (a) Mean of each sample for the classification dataset and (b) variance of each sample in the dataset (150 samples).....	52
Figure 45. (a) Spectral Energy spread for the classification dataset and (b) Temporal Energy Spread in the dataset (150 samples).....	53
Figure 46. Architecture of the Multi-Layer Perceptron Network used.....	55
Figure 47. Training Loss for the MLP Network	56
Figure 48. Training MAE for the MLP Network.....	56
Figure 49. Training Accuracy for the MLP Network	57
Figure 50. Prediction Results for the MLP Network	58
Figure 51. Normalized similarity number $n_i^m(r)$ vs number of pits	62
Figure 52. Normalized similarity number $n_i^m(r)$ vs defect depth.....	63
Figure 53. (a) HT responses for increasing number of pits and (b) a zoomed-in version of Peak II	64
Figure 54. (a) HT responses for increasing defect depth and (b) a zoomed-in version of Peak II	65
Figure 55. Multi-Layer Perceptron network for defect response classification.....	65
Figure 56. (a) Prediction results for the MLP network and (b) confusion plot for the predicted results	66
Figure 57. (a) Pure raw signal and (b) signal with additive white Gaussian noise (AWGN) added at SNR 5	66
Figure 58. (a) Prediction results for the MLP network trained with a dataset with SNR 20 (b) confusion plot for the predicted results.....	67
Figure 59. (a) Prediction results for the MLP network trained with a dataset with SNR 10 (b) confusion plot for the predicted results.....	68
Figure 60. (a) Prediction results for the MLP network trained with a dataset with SNR 5 (b) confusion plot for the predicted results.....	68
Figure 61. Architecture of the proposed 1D-CNN.....	69
Figure 62. Sample 11 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	70

Figure 63. Sample 12 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	71
Figure 64. Sample 13 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	71
Figure 65. Sample 14 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	72
Figure 66. Sample 15 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	72
Figure 67. Sample 16 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	73
Figure 68. Sample 17 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	73
Figure 69. Sample 18 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	74
Figure 70. Sample 19 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	74
Figure 71. Sample 20 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.	75
Figure 72. Scatter plots of burst pressure (P_b) vs. selected quantities	82
Figure 73. A quarter of a corroded pipe modeled in ABAQUS	83
Figure 74. Scatter plots of selected quantities vs. yield stress (σ_y).....	85
Figure 75. Comparison of residual and MSE of each model for three levels of σ_y	89
Figure 76. Comparison of MSE of each model for three levels of D/t , d/t , and l^2/Dt	91
Figure 77. Scatter plots of (a) σ_y vs. grade and (b) σ_u vs. grade.....	93
Figure 78. Scatter plots of the failure pressure predicted by the proposed models vs. the observed pressure	96
Figure 79. Comparison of residual and MSE of the proposed models with the existing models for three levels of σ_u	97
Figure 80. Scatter plots of the prediction of proposed model and best existing model vs. the validation data (20% of data).....	98
Figure 81. Sensitivity of the three selected quantities on pressure prediction.....	100
Figure 82. Configuration of two defects	103
Figure 83. Combining interacting defects.....	108
Figure 84. Scatter plots of burst pressure (P_b) vs. selected quantities	110
Figure 85. A corroded pipe modeled in ABAQUS.....	111
Figure 86. Scatter plots of selected quantities vs. yield stress (σ_y).....	113
Figure 87. Comparison of the existing interaction rules using P_{CD}	115
Figure 88. Comparison of existing failure pressure models	116
Figure 89. Comparison of the proposed rule (PR) with three best existing interaction rules.....	119
Figure 90. Comparison of the proposed model (PM) with the MTI method and two Level-1 assessment methods	121
Figure 91. Illustration of the FAD	126
Figure 92. Scatter plots of burst pressure (P_b) vs. selected quantities	129

Figure 93. Crack propagation in a pipe modeled in Abaqus.....	131
Figure 94. Scatter plots of selected quantities vs. yield stress (σ_y).....	133
Figure 95. Comparison of existing failure pressure models	136
Figure 96. Comparison of P_b/P_a using the modified Ln-Sec model (Mod Ln-sec) and the proposed model (PM)	138
Figure 97. Sensitivity of the four selected quantities on pressure prediction	139
Figure 98. Defect depth-dependent reliability index based on the proposed models and best existing models for pipelines with single corrosion defect.....	142
Figure 99. Absolute values of importance measures based on the proposed models for pipelines with single corrosion defect assuming $l = 100\text{mm}$	143
Figure 100. Defect depth-dependent reliability index calculated with and without the model error for the proposed models and best existing models for pipelines with single corrosion defect...	145
Figure 101. Reliability index versus operating time based on different interaction rules	147
Figure 102. Time evolution of a/d_w , P_b , and P_f calculated for Case Study III	151
Figure 103. Expected total cost versus inspection interval considering $P_f(t)$ evaluated by the proposed model (PM) and the modified Ln-Sec for pipelines with single crack-like defect	151

Table 1. Properties of Steel AISI 4340	31
Table 2. Parameters for CSamp-En	62
Table 3. Comparison of 24 existing pressure failure prediction models	79
Table 4. Data range of six important quantities.....	80
Table 5. Experimental testing cases selected from literature for validating the FE models.....	86
Table 6. New FE models cases	87
Table 7. Comparison of the best models of subsets.....	94
Table 8. Final selected model formula and model parameter statistics	96
Table 9. Existing interaction rules	104
Table 10. Existing capacity pressure prediction models.....	105
Table 11. Summarized procedure for DNV RP-F101 for interacting defects (Part B) and MTI method.....	106
Table 12. Summary of the database collected from literature	109
Table 13. Selected cases for FEM validation.....	112
Table 14. Summary of the database generated from FE	114
Table 15. Number of data points.....	114
Table 16. Variables and model parameter statistics for the final interaction prediction model .	118
Table 17. Variables and model parameter statistics for the correction factor	121
Table 18. Summary of the database collected from literature	128
Table 19. Experimental testing cases selected for calibrating <i>Maxpe</i> and <i>G_c</i> and calibrated results	132
Table 20. New FE models cases	134
Table 21. Conditions for using the CorLAS , API 579 and BS7910 models in this study	135
Table 22. Variables and model parameter statistics for the correction factor	137
Table 23. Distribution parameters of random variables used in Case Study I.....	141
Table 24. Distribution parameters of random variables used in Case Study II	146
Table 25. Distribution parameters of random variables used in Case Study III	150

List of Acronyms

AIC	Akaike Information Criterion
AWGN	Additive white Gaussian noise
BIC	Bayesian Information Criterion
CNN	Convolutional Neural Network
CSamp-En	Cross sampling entropy
DI	Damage index
DNV	Det Norske Veritas
EC	Eddy current
ECA	Eddy current array
ECT	Eddy current testing
EMAT	Electro-magnetic acoustical transducer
FAD	Failure assessment diagram
FCN	Fully connected network
FE	Finite element
FEM	Finite element modeling
FN	False negative
FORM	First Order Reliability Method
FP	False positive
HT	Hilbert transform
IFM	Infinite microscopy
IFM	Infinite microscopy
ILI	In-line inspection
KH	Kere and Huang
MAE	Mean absolute error
MCEA	Multi cross entropy analysis
MFL	Magnetic flux leakage
ML	Machine learning
MLP	Multi-layer perception
Mod Ln-Sec	Modified Ln-Sec model
MSE	Mean squared error
MTI	Mixed-type interaction
NDE	Nondestructive evaluation
PR	Proposed rule
Proposed model	Proposed model
QUAD	Quadrilateral
RFEC	Remote field eddy current
SampEn	Sampling entropy
SCC	Stress cracking corrosion
SGD	Stochastic gradient descent
SH	Shear horizontal
SNR	Signal to noise ratio

SORM	Second Order Reliability Method
TN	True negative
TP	True positive
UGW	Ultrasonic guided waves
UT	Ultrasonic testing
XFEM	Extended finite element method

1 Executive Summary

To address DOT's pressing need for safety and integrity maintenance of the existing pipeline infrastructure in the U.S., the proposed project focuses on multi-modal NDE and probabilistic performance evaluation of aging pipelines under interactive threats. This study will utilize experimental testing and numerical analysis to generate more realistic defect shapes and colony profiles, which will be used for characterization and validation of interactive defect NDE. In addition, probabilistic models of failure pressure of a pipeline containing corrosion and cracking-like defects will be developed, achieving predictions that are unbiased with reduced variability and considering defect interaction.

The technical impact of the proposed research can be summarized in four aspects: (1) The proposed multi-modal NDE framework enables the missing capability to assess interactive anomalies with the integration of lab-, field- and simulation-environment validation. (2) Various sources of uncertainties are quantified and appropriately propagated to risk assessment through probabilistic characterizing defect profiles in NDE, probabilistically modeling time-evolution of defect profile propagation, development of probabilistic capacity model considering interactive anomalies, and reliability analysis. (3) The developed probabilistic capacity model remove bias and improve the accuracy of the deterministic models, complement the deterministic models with characteristics of defect profiles, and preserve the simplicity of the deterministic models so as to enable the practical application of the proposed probabilistic models. (4) The corrosion defect interaction impact are probabilistically assessed, which is suitable for risk assessment. Overall, the results of reliable performance predictions generated from this research enable optimum monitoring/inspection, maintenance scheduling/methods, repair strategies/methods, and financial resource allocations and forecasting.

The proposed project result in (1) a better understanding of the characterization of interactive anomalies in isolated and colony profiles using NDE and their impacts on the residual strength of a pipeline; (2) industry ready probabilistic prediction models for failure pressure of pipelines containing interactive anomalies, providing predictions that are unbiased with reduced variability; and (3) better knowledge of the propagation and quantification of prevailing uncertainties in prediction models for the quantitative risk management of pipelines.

1.1 Summary of Accomplishments

Peer-reviewed journal publications

- Hamilton, C., Miles, Z., Zhou, Q., Huang, Q., Chakrapani, S, and Deng, Y. (2023). “Multi-modal NDE and data fusion for interactive corrosion detection and characterization”, *NDT&E International* (in preparation).
- Kere, K.J., and Huang, Q. (2023). “Probabilistic models of defect interaction identification and failure pressure for pipelines with colony of corrosion defects”, *International Journal of Pressure Vessels and Piping* (under review).
- Kere, K.J., and Huang, Q. (2023). Probabilistic burst pressure prediction model for pipelines with single crack-like defect,” *International Journal of Pressure Vessels and Piping*, 105084.
- Kere, K.J., and Huang, Q. (2023). Expected life-cycle cost evaluation for deteriorating pipelines,” *Reliability Engineering and System Safety* (under revision).
- Kere, K.J., and Huang, Q. (2022). “Probabilistic Burst Pressure Prediction for Corroded Pipelines,” *Journal of Pressure Vessel Technology*, 104656.

Conference proceedings

- Kere, K.J., and Huang, Q. (2023). “Development of probabilistic models of defect interaction identification and failure pressure for pipelines with colony of corrosion defects,” *14th International Conference on Applications of Statistics and Probability in Civil Engineering*, AMPP Annual Conference & Expo 2023.
- Kere, K.J., and Huang, Q. (2023). “Risk management strategies and probabilistic failure pressure model development for pipelines with crack-like defect,” *14th International Conference on Applications of Statistics and Probability in Civil Engineering*, ICASP14, Dublin Ireland.
- Rautela, M., Gopalakrishnan, S., Gopalakrishnan, K., and Deng, Y. (2020). “Ultrasonic guided waves-based identification of elastic properties using 1d-convolutional neural networks”, *2020 IEEE International Conference on Prognostics and Health Management (ICPHM)*, p. 1-7. Citation: 21 (as of 07/2023)

Oral and Poster Presentations

- Kere, K.J., and **Huang, Q.** “Risk management strategies and probabilistic failure pressure model development for pipelines with crack-like defect,” *14th International Conference on*

Applications of Statistics and Probability in Civil Engineering, ICASP14, Dublin Ireland, 2023.

- Rautela, M., Gopalakrishnan, S., Gopalakrishnan, K., and Deng, Y. (2020). “Ultrasonic guided waves-based identification of elastic properties using 1d-convolutional neural networks”, *2020 IEEE International Conference on Prognostics and Health Management (ICPHM)*, p. 1-7. Citation: 21 (as of 07/2023)
- Kere, K.J., and **Huang, Q.** “Development of probabilistic models of defect interaction identification and failure pressure for pipelines with colony of corrosion defects,” *14th International Conference on Applications of Statistics and Probability in Civil Engineering*, AMPP Annual Conference & Expo 2023.
- Kere, K.J. and Huang, Q. “Development of probabilistic models of defect interaction identification and failure pressure for pipelines with colony of corrosion defects,” ASCE UESI Pipelines 2022, Indianapolis, Indiana, July 31 – August 3, 2022.
- Kere, K.J. “Probabilistic models development and risk management strategies for pipeline with anomalies”, 2022 Forward Thinking Research Symposium, Marquette University.
- Kere, K.J., and Huang, Q. “Multi-modal NDE Assisted Probabilistic Pipeline Performance Evaluation under Interactive Anomalies,” 2021 PHMSA R&D Forum.
- 2020 R&D Forum

Student Contributors

- Kere, Jules (PhD student, Marquette University)
- Adam, Puchalski (undergraduate student, Marquette University)
- Zichen Ling (PhD student, The University of Akron)
- Kuan-Chen Huang (MS student, The University of Akron)
- Hamilton, Ciaron (PhD student, MSU)
- Miles, Zeb (PhD student, MSU)
- Karthik Gopalakrishnan (MS student, MSU)

2 Introduction/Background and Objectives

Oil and gas pipelines are a critical part of the infrastructure of modern society. The U.S. has about 3 million miles of gas and liquid pipelines (including more than 400,000 miles of transmission pipelines), subjected to various potential threats during their service lives. In the past three decades, qualitative risk management has been widely used in industry practice for supporting cost-effective decisions to achieve specific acceptable levels of safety. Figure 1 shows a general process of the quantitative risk management for a pipeline, consisting of exposure, risk quantification, and decision-making. Once the potential threats are identified, the risk associated with those threats need to be assessed so that appropriate actions (such as mitigation, prioritizing of maintenance, repair and replacement) can be taken. Risk is typically defined as the product of the probability of failure and the consequence of failure.

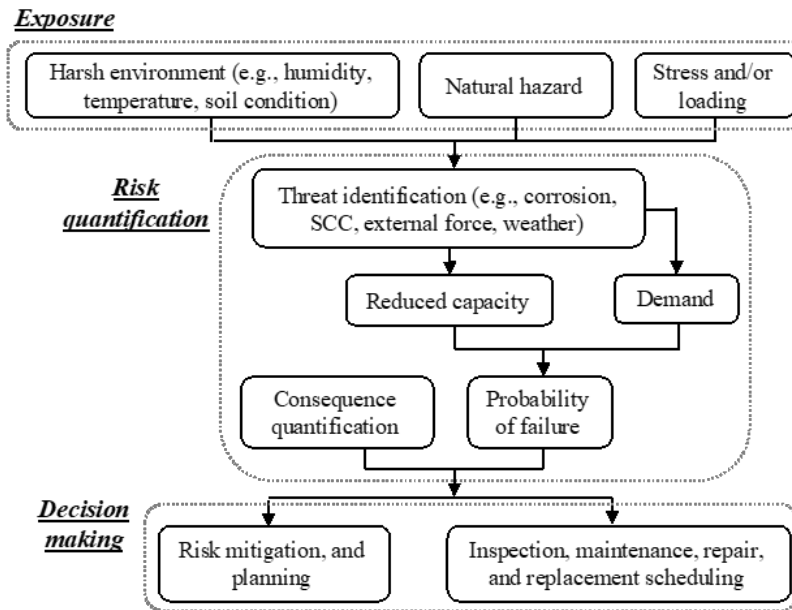


Figure 1. Quantitative risk management for pipeline.

Based on the degree of impact for a consequence, a pressurized pipeline can fail in two distinctive failure modes: a small leak or a burst. In particular, a burst occurs when internal pressure exceeds the pressure resistance of the pipeline, and it usually leads to significant safety and environmental consequences. The capacity to resist internal pressure (or *failure pressure*) is reduced by the damage defects resulting from accidental impacts or material degradation such as metal corrosion and cracks. When damage defects are considered as a threat, nondestructive evaluation (NDE) technologies are able to detect the location and geometry of a damage defect with a high

degree of accuracy, such as magnetic flux leakage (MFL), eddy current (EC) including single frequency, multiple frequency and pulsed excitations, ultrasonic testing (UT). In recent years, technologies like electro-magnetic acoustical transducer (EMAT) in-line inspection (ILI) were developed that can capture many crack-like features and address undiggable challenges through research projects sponsored by PHMSA, however, the reliability and accuracy of using NDE and/or ILI for crack detection still need to be continuously improving. One key reason is that state-of-the-art technologies are limited in identifying and characterizing interactive anomalies.

Recent PHMSA studies have also confirmed that many pipe failures are not resulted from a single type of threat but threat interactions, including interactions of resident conditions with changing operations or environment. *There is a major technical gap regarding characterization of the interactive anomalies and reliability assessment of pipeline under such anomalies.* Therefore, to obtain accurate risk assessment, one needs a thorough understanding of the time-dependent physical characteristics of interacting damage from advancements in NDE methodologies, the level of operation loading demand, the probabilistic capacity assessment considering time-evolution of anomalies, and quantification of all relevant uncertainties.

2.1 Objectives

The goal of this proposed study is to develop a probabilistic pipeline performance evaluation framework based on multi-modal NDE assisted by physical and mechanical modeling under interactive anomalies. This study utilizes experimental testing and numerical analysis to generate more realistic defect shapes and colony profiles, which will be used for characterization of interactive defects and validation of NDE. Meanwhile, the identified defect profile are used for the probabilistic defect time-evolution model development, which is crucial for reliability evaluation of pipeline performance under interactive defects. In addition, probabilistic models of failure pressure of a pipeline containing corrosion and cracking-like defects are developed, achieving predictions that are unbiased with reduced variability and considering defect interaction.

Specific technical objectives are as follows:

- Objective 1: Generate realistic corrosion and cracking defect profiles through laboratory testing and electrochemical simulation;
- Objective 2: Establish an expanded NDE framework for interactive anomalies by probabilistic characterization of defect profiles;
- Objective 3: Develop probabilistic failure pressure prediction models incorporating defect

interaction;

- Objective 4: Investigate the impact of various physical quantities and uncertainty sources on pipeline reliability.

2.2 Justification of Scope Adjustment

For the experimental testing, the budget and time restrained us to explore the possibility to generate crack corrosion which is found to be much harder than general corrosion under lab conditions. Consequently, simulation of crack corrosion was not able to be conducted as no lab results are available for calibration. Here are the specific challenges that we encountered in the lab testing:

- In the initial experimental testing, generating a general corrosion profile took almost 1.5 years using flat samples under B117 and G85 environment. However, we found that the corrosion defect results cannot be used directly for probabilistic analysis because there are significant sample variations, and the corrosion depth profile unfortunately did not present a clear trend. The flat sample testing took around 1.5 years, and the samples were lost in the shipping transportation to MSU for NDE characterization.
- Then, we conducted another corrosion environmental test under B117 exposure by using pre-damaged samples where a designed defect was introduced through grooving the samples. The B117 data are reasonable and can be used for probabilistic analysis and NDE testing. However, the pre-damaged samples testing took another 1.5 years.
- Generating general corrosion defects alone took about 3 years, which exhausted all the budget that was budgeted for the 1.5-year Task 1; we completed the experimental part for corrosion depths data within the project time frame with the supplement of internal funding.

For the NDE tasks, the NDE framework was not able to be tested in a field environment due to the difficulty of obtaining such data. However, MSU has focused their efforts on developing methodologies to characterize interacted corrosion defects. In addition, MSU also has tested the developed methodologies on the samples with general corrosion sent by UAkron. The step of using NDE data for reliability analysis was not complete due to time and budget constraints but it will be completed using other internal funding with the continued collaboration of the three universities.

For probabilistic capacity model and reliability analysis, Marquette Research team has conducted a comprehensive work for pipeline with three scenarios: isolated corrosion defect,

colony of corrosion defects, and isolated crack defect. We encountered research challenges regarding the other two scenarios (i.e., colony of crack defect and colony of corrosion and crack-like defects), which cannot be addressed during the project period. First of all, there is very limited existing data in the literature for colony crack-like defect or colony of corrosion and crack-like defects. It is worth mentioning that one study related to colony of crack-like defect has been experimentally conducted by a PRCI project that was completed in 2020 (<https://www.prci.org/192422.aspx>); however, it contains only four burst tests. In addition to the lack of existing data, the plastic properties (or J-R curves) for the pipeline that are needed for numerical modeling are not reported in the literature, which makes the modeling validation impossible for these two scenarios. In summary, it needs one or two separate research projects to systematically investigate the last two scenarios as the experimental data is extremely limited.

3 Objective 1: Lab Testing of Generating Realistic Corrosion Defect

3.1 Experimental Program

This objective is to generate realistic corrosion profiles through environmental exposure testing. The defect shapes and colony profiles will be used for NDE and for the probabilistic defect time-evolution model development.

The testing metal is a ground low-carbon steel with a similar composition to the API series pipeline metals. One type of metal sample was the flat sheet with the size of 3" × 3" × 3/32". The second type of metal sample was the same flat sheet that was punched to generate a pre-damaged area on the surface. The punched scratch on the surface was around 0.63" long, 0.06" wide, and 0.010" deep. The exact depth of the scratch was measured by infinite microscopy (IFM) measurement before and after exposure testing.

One testing condition was the continuous salt spray of 5 wt.% NaCl fog following ASTM B117 salt spray testing protocol. The second testing environment is according to the ASTM G85 standard where the samples were in exposure to a mixed solution consisting of 0.35 wt.% aluminum sulfate and 0.05 wt.% NaCl.

After certain periods of time, the testing samples were removed from the environmental chamber for IFM characterization. The surface photos of the testing samples after exposure testing were included in Appendix A.

3.2 Results and Discussion

3.2.1 Flat sample results

The infinite microscopy images for the testing flat samples from initial immersion to 8-week exposure in B117 environment are shown in Figure 2. The surface area of the metal under the exposure is 3" × 3" (demonstrated as the blue square in the 1st row of the figure), while the surface area for the infinite microscopy scanning is 3mm × 3mm (demonstrated as the red square in the 1st row of the figure). Five different locations on the testing metal were chosen for the infinite microscopy scanning, which is center, left top, left bottom, right top, and right bottom of the surface. Based on the change in the topography of the surface, the evolution of a corrosion profile can be observed. For example, some small corrosion spots can be detected on the right bottom of the surface after 8-week exposure.

Besides topography, two quantities can be obtained from the infinite microscopy scanning: the average depth (d_{avg}) and the maximum depth (d_{max}) of the scanning area. To evaluate the depth of the whole surface area, the sample mean and sample standard deviation of d_{avg} and d_{max} from the five scanning locations were calculated and shown in Figure 3. The average depth and the maximum depth for duplicated samples are also included in Figure 3. In general, the means of d_{avg} and d_{max} increased with time during the 24-week exposure time in the B117 condition. The increase of d_{max} demonstrates the generation of corrosion depth under the corrosive environment.

The same flat samples were also exposed in the environmental chamber following the ASTM G85 testing protocol and investigated by infinite microscopy characterization for the average depth and the maximum depth. Figure 4 shows the sample mean and sample standard deviation of d_{avg} and d_{max} from the five scanning locations of all the exposed samples during the 24 weeks of G85 exposure. The average depth and the maximum depth for duplicated samples are also included in Figure 4. As expected, the samples showed a higher corrosion depth under G85 immersion, which was a harsher environment than the B117 exposure.

The fluctuations of the depth values are due to the sample variations because different samples were removed from the environmental chamber and tested at each testing period. After the pre-analysis of the IFM results of the flat samples, the big fluctuation brings large errors in probabilistic analysis, which cannot be further used for the analysis.

In addition, all the flat testing samples were shipped to the Co-PI Dr. Deng at Michigan State University for NDE characterization at the end of March 2021, but all these samples were lost in transportation by USPS. Therefore, a second testing sample with the pre-damaged surface

was designed for exposure testing.

3.2.2 Pre-damaged sample results

The infinite microscopy characterizes the depth of the pre-damaged surface area of the testing sample. The details of IFM testing are described in Appendix A. One sample was continuously placed in the chamber, which means this sample was put back to the exposure after IFM characterization periodically, and its depth is shown in Figure 5. The other set of samples was initially put in the exposure and removed at a certain time, and their depths are shown in Figure 6. The depth change is defined below:

$$\text{Depth change} = \text{depth (time of testing)} - \text{depth (initial)} \quad (3.1)$$

Figure 7 shows the mean and standard deviations of the depth changes of all the samples. As shown in Figure 7, the change of depth presents a significant increase in the B117 exposure environment during the 36 weeks of exposure. The duplicated samples show small variations.

Similarly, for G85 exposure testing, the depth of the continuously exposed sample is shown in Figure 8 and the other set of the testing samples is shown in Figure 9. The depth change of all the testing samples under the G85 exposure is shown in Figure 10. The depth change is less significant during the 36 weeks of exposure under the G85 environment, and the sample variations are large.

The results of pre-damaged samples in B117 exposure demonstrate a good depth change and the sample variations are small. These samples after testing were sent to the Co-PI Dr. Deng at Michigan State University for NDE characterization in the fall of 2022.

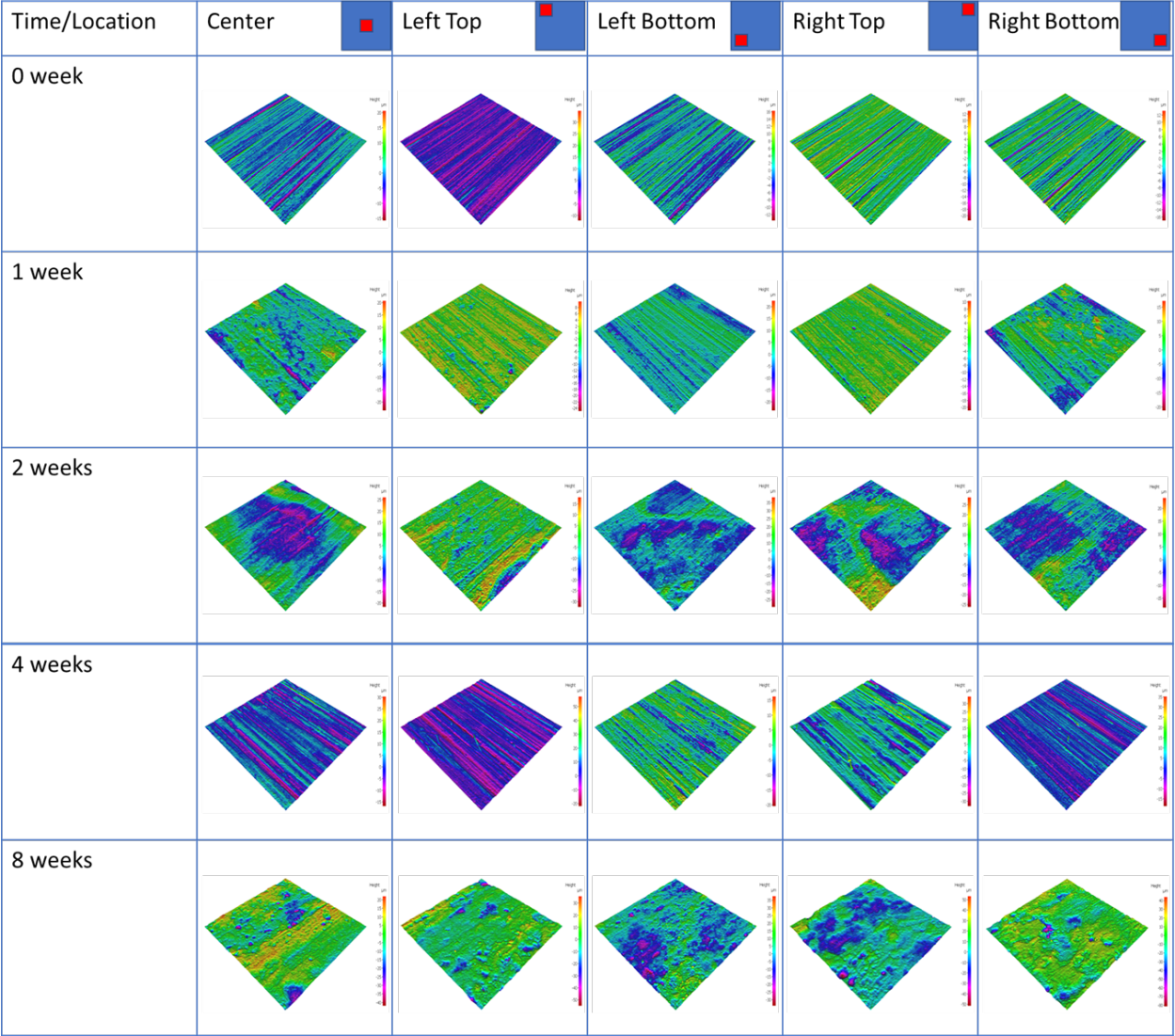


Figure 2. Infinite microscopy images of the flat testing samples during B117 exposure testing

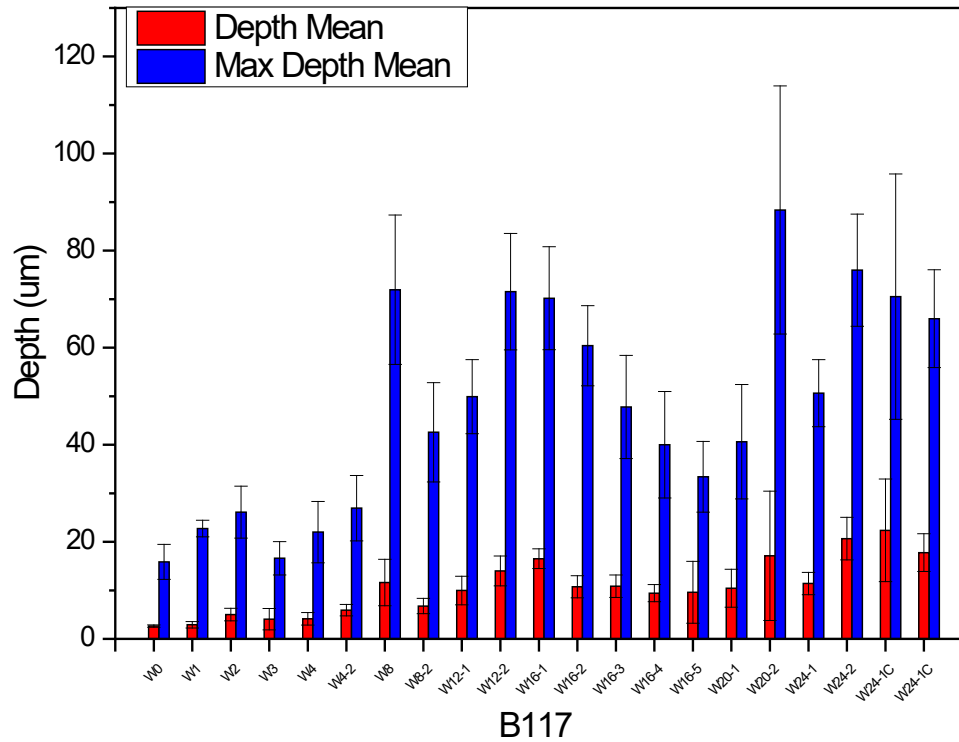


Figure 3. The sample mean and sample standard deviation of the average depth and the maximum depth of the scanning area of all the flat samples during B117 exposure testing.

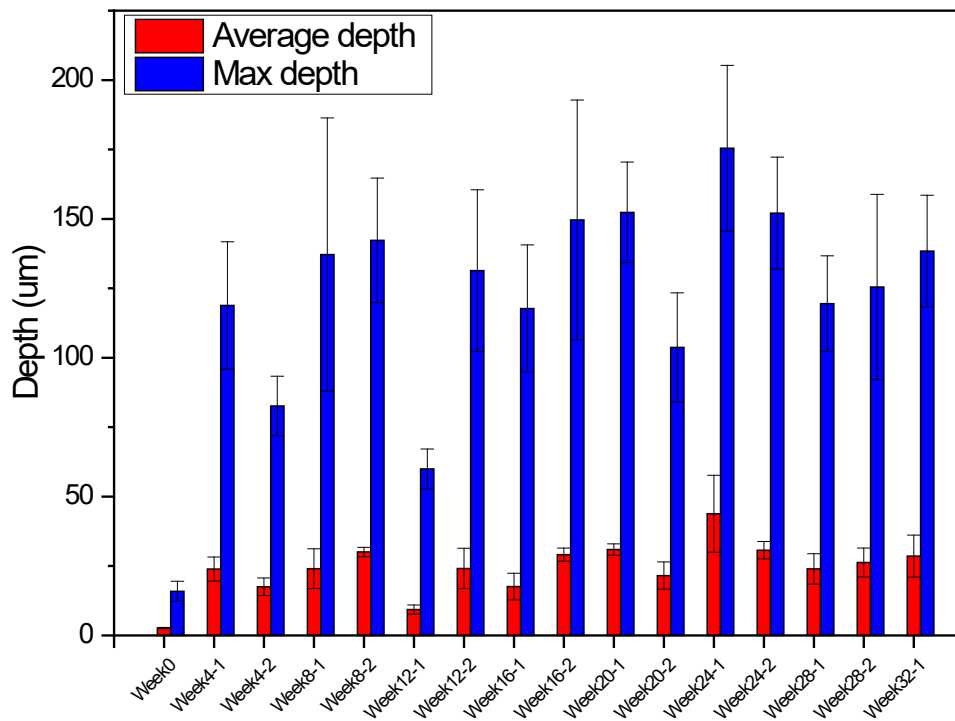


Figure 4. The sample mean and sample standard deviation of the average depth and the maximum depth of the scanning area of all the flat samples during G85 exposure testing.

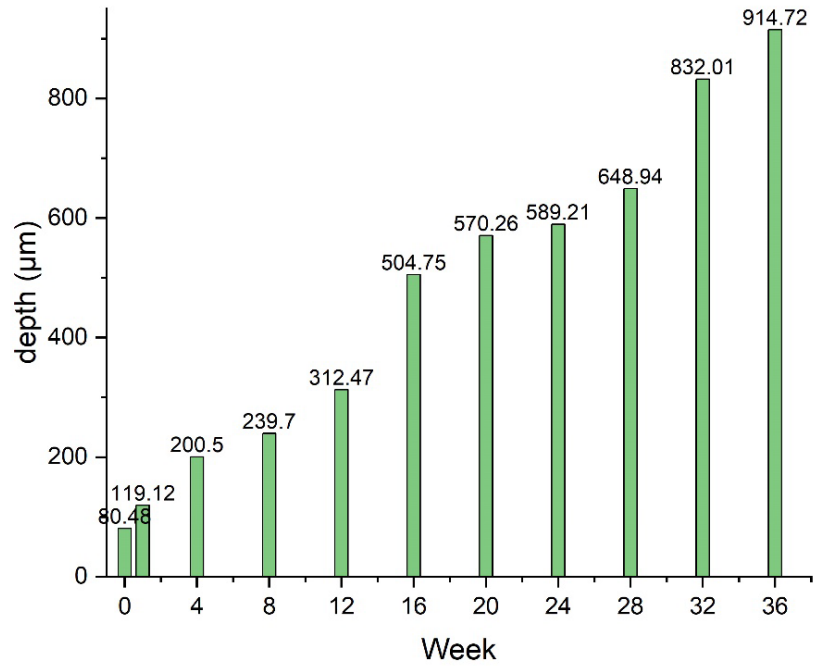


Figure 5. The depth of the pre-damaged testing sample continuously under B117 exposure.

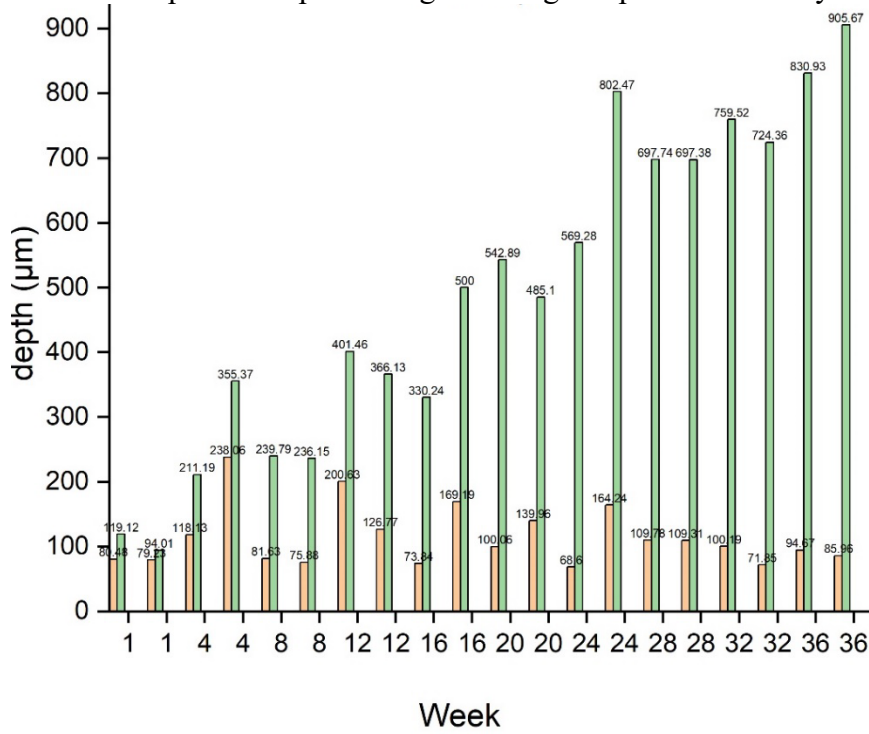


Figure 6. The depth of the pre-damaged testing samples under B117 exposure. Orange: initial depth; green: the depth at the time of testing.

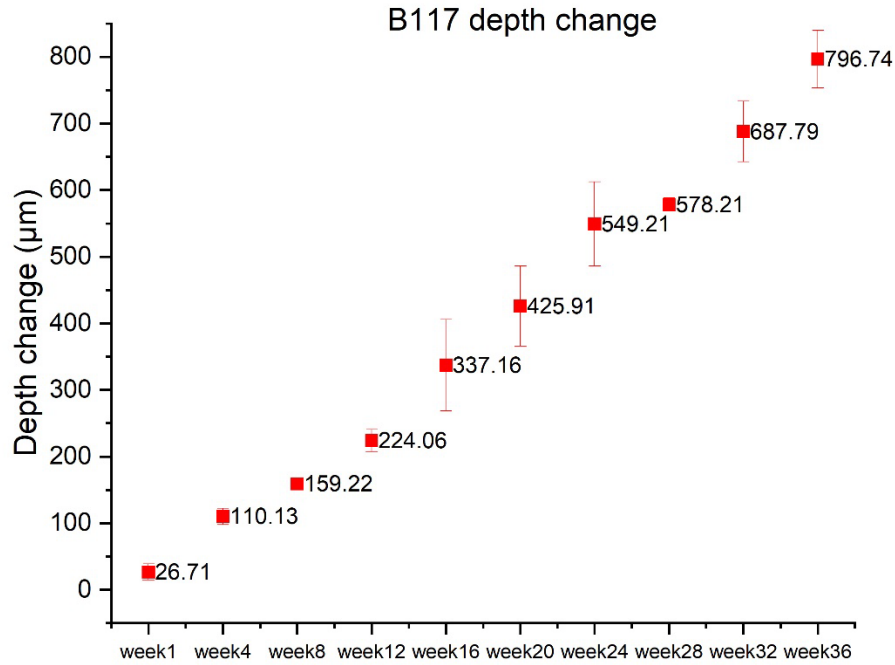


Figure 7. The change of the depth of the pre-damaged testing samples under B117 exposure.

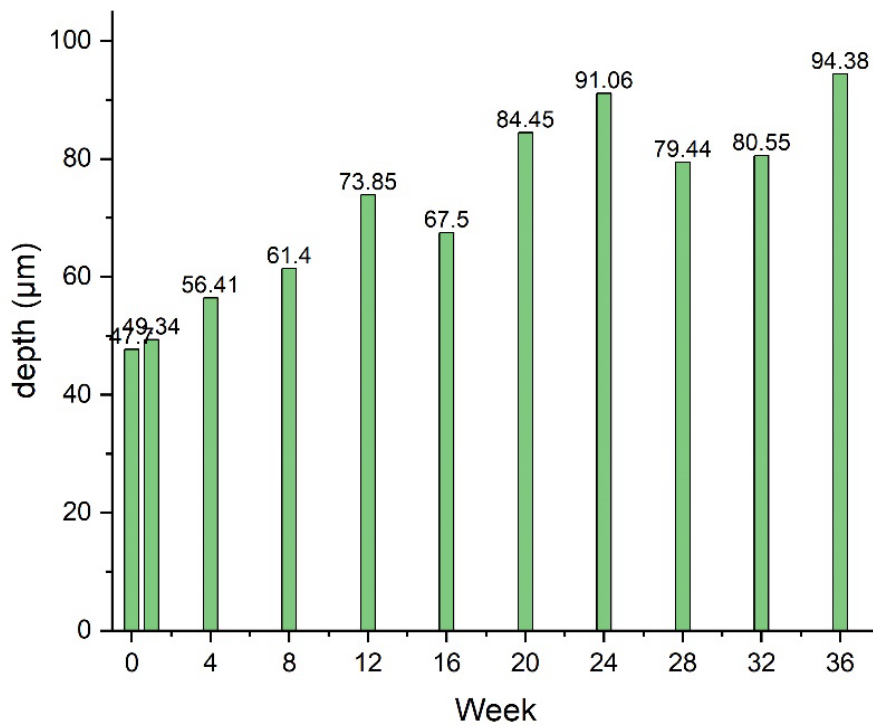


Figure 8. The depth of the pre-damaged testing sample continuously under G85 exposure.

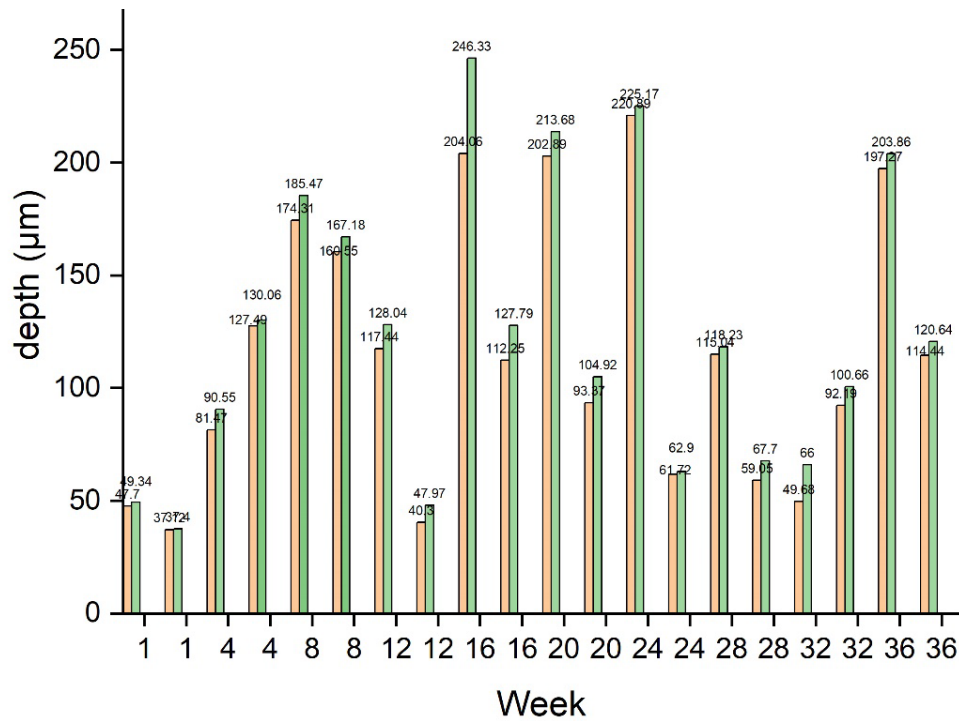


Figure 9. The depth of the pre-damaged testing samples under G85 exposure. Orange: initial depth; green: the depth at the time of testing.

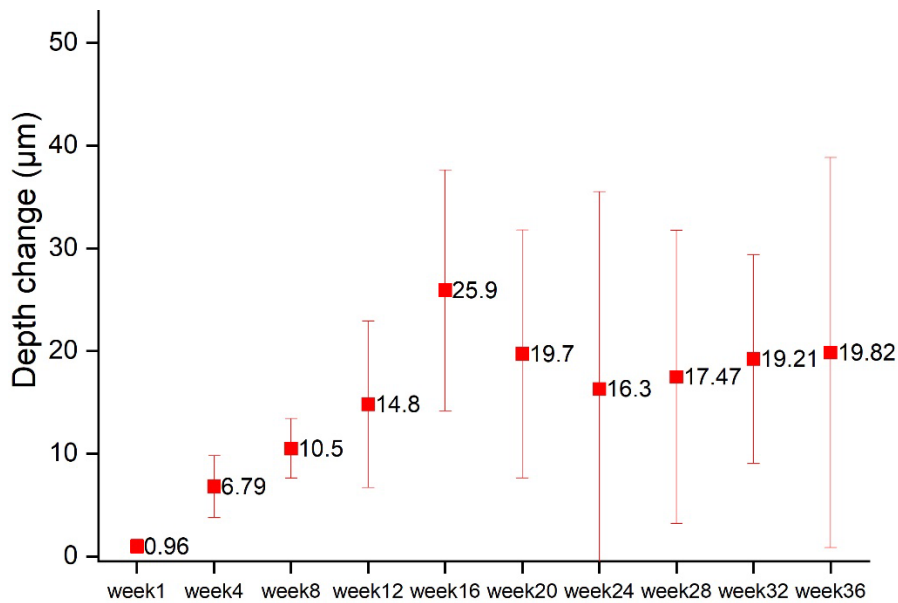


Figure 10. The change of the depth of the pre-damaged testing samples under G85 exposure.

4 Objective 2: NDE Framework

4.1 Eddy Current Array Methods for Interactive Corrosion Detection and Characterization

Interactive defects detection and characterization in metallic pipes is one of the major challenges identified for pipeline integrity assessment. The group here previously has developed/is developing novel NDE and data processing methods for pipeline applications, including internal corrosion inspection using optical structured light 3D reconstruction and rendering techniques that significantly improves the damage detectability, and stress cracking corrosion (SCC) detection using multi-frequency electromagnetic techniques, remote field eddy current (RFEC) techniques, etc. assisted by machine learning (ML). While there are tremendous successes in these techniques, which work well for exposed pipes or “in-the-ditch NDE”, only Shear Horizontal (SH) guided wave testing has been proven to work in NDE of buried pipelines that poses a big challenge in field-testing to understand realistic interacting threats environment. In this task, the MSU NDE team develops a multi-modal electromagnetic and ultrasonic framework including Eddy Current Array (ECA), EMAT for generation of SH waves (low frequency-50kHz to 500 kHz), localized Rayleigh wave measurement using EMATs and air coupled transducers, and contact ultrasonic measurement for validation of guided wave results for better characterizing the identified interactive anomalies, as well as leveraging other techniques being developed by the group through the other successful programs sponsored by PHMSA. Defect localization and material characterization have always been a challenge for guided waves inspection in this community; and it is worth noted since SH waves have very little out-of-plane leakage, their energy is confined within the pipe walls and they can propagate for long distances. Therefore, any local changes to thickness or material degradation (loss in stiffness and density) can be detected using SH waves, which makes it a perfect candidate for the proposed corrosion/SCC/fatigue defects interaction study. Expanding from the ongoing PHMSA project, introduction of SH waves modality and dedicated signal processing algorithms for analyzing the interactive-damage-feature-encoded data is crucial for the success of the proposed work.

4.1.1 ECA theory and parameter selection

Eddy current testing (ECT) is a low-cost and robust method for nondestructive evaluation for various inspection purposes and matches well for corrosion detection which contains small defect depths. Rugged ECT sensors may be designed at high frequencies required from sub-mm damages.

As the coils will be small to meet high frequencies, they may also be duplicated as an ECA for faster scan times. This report demonstrates the concepts for designing or selecting an eddy current probe, and the procedure for scanning with a gantry system.

The input to an ECT coil is a signal from a function generator with an input voltage v_{in} and frequency f . The output will be response voltage v_{out} affected by complex impedance and induced current density on the sample under test. Vital information for how to conduct ECT depends on the size of defect and material used. This is relevant to skin penetration, which determines the depth of the response, which is defined by:

$$\delta = \frac{1}{\sqrt{f\sigma\pi\mu}} \quad (4.1)$$

With δ being the skin penetration depth, f being the frequency, σ being electrical conductivity in S/m , and μ being magnetic permeability in H/m , with $\mu = \mu_r * \mu_0$ where μ_r is realative permeability (unitless) and μ_0 being the permeability of free space with $\mu_0 = 0.4\pi * 10^{-6}H/m$. By defining the thickness for skin depth δ for materials determining σ and μ , a frequency f may be selected based on coil properties. From depth measurements using infinite focus microscope (IFM) measurements, the range of damage is between about $100\mu m$ and $900\mu m$. For steel, conductivity has been measured to be $4.68 \times 10^6 S/m$ on steel and $0.75 \times 10^6 S/m$ within the corroded region, while relative permeability was measured around 60 for steel and 4 for the corroded region [1]. Because of the differences between the electromagnetic properties of steel and rust, the response of the eddy current density will also change, which change may be measured through v_{out} . To match the depth of corrosion at $900\mu m$, a frequency around $104kHz$ is desired. Keep in mind there are several variables that keep this selection from being perfect. For example, how steels are processed will vary its electromagnetic properties. Another factor is lift-off between a coil in air versus the sample, which conductivity is suggested to be approximately zero S/m and relative permeability of 1 [2]. If the lift-off is too high, this will decrease the response of the signal which is not wanted.

There are other important prospects to ECT scanning, including the usage of a gantry for scanning and array probes for faster scanning. A gantry may hold onto ECT probe provides accurate positional information via encoders with respect to throughput ECT data. Gantries may also read in commands for program mable scanning. Raster movement patterns are used, which will provide a 2D image of the scan. Since coils used for ECT testing may have small diameters

for high frequency testing, required for detection of corrosion with small damages, they can be replicated in an array pattern to decrease scan times. For example, a single coil running a raster scan of $X = 30mm$ and $Y = 30mm$ at a desired $1mm$ resolution along the shifting axis X , with scanning axis Y being dependent on the data acquisition frequency and gantry velocity. A single probe requires 30 shifts along X to match the required resolution. If two coils c_1 and c_2 are implemented, strategically placed $1mm$ away from each other, then c_1 may skip c_2 for every shift to avoid redundancy, requiring 15 shifts which effectively decreases scan time by half. The downside to array probes is that each coil requires calibration in terms of gains to ensure each coil outputs similar output voltages.

4.1.2 Procedure and results

Ten steel samples containing weekly ASTM b117 standard corrosion were scanning, shown in Figure 11. Sample 1 starts at 1 week's worth of corrosion, which the consecutive samples are corroded for 4 weeks afterwards until week 36. Each sample initially contained a defect about $15 \times 5mm$ in area and approximately $0.1mm$ in depth. An I-Flex ECA probe with an operating frequency range between $100 - 800kHz$ and 32 channels was used alongside an Ectane 2 testing instrument and Magnifi eddy current data acquisition and processing software. MATLAB was used for further post processing. Each channel is connected to a coil, with each coil being $2mm$ in diameter. The array contains 2 columns of 16 rows probes along the X axis with a shift of $1mm$ between each column. This effectively gives the coverage at $34mm$ with a resolution of $1mm$ from one swipe. However, because of lift off variation between the two columns, combining both results give unwanted results due to lift-off variation. It was decided to split the data between the two columns to obtain two different images, shown in Figure 12 and Figure 13, with a $1mm$ shift along the Y axis between the two sets. This mitigates any tilt along the Y axis between columns during calibration and scanning, leaving only tilt along each individual column only the X axis.

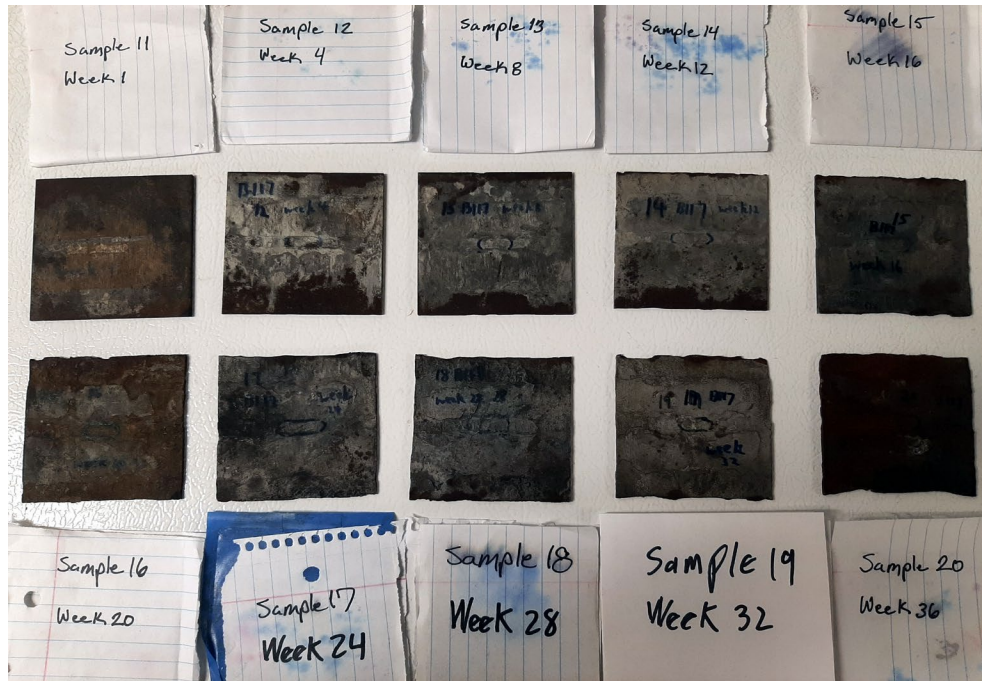


Figure 11. Corroded samples and their respective labels.

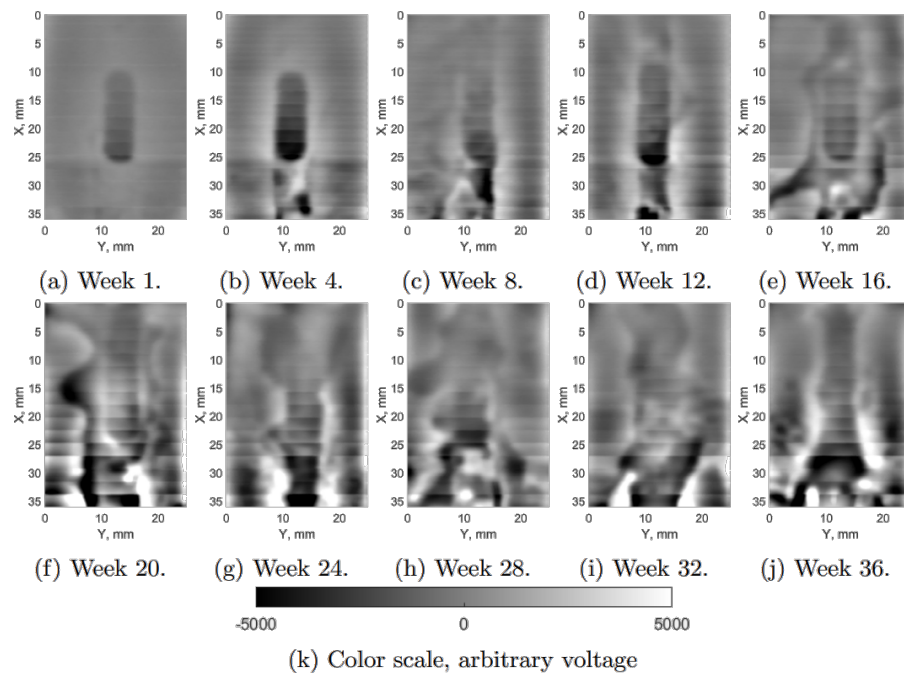


Figure 12. 2D post-processed data from real voltage component on corroded samples using the left column of coils.

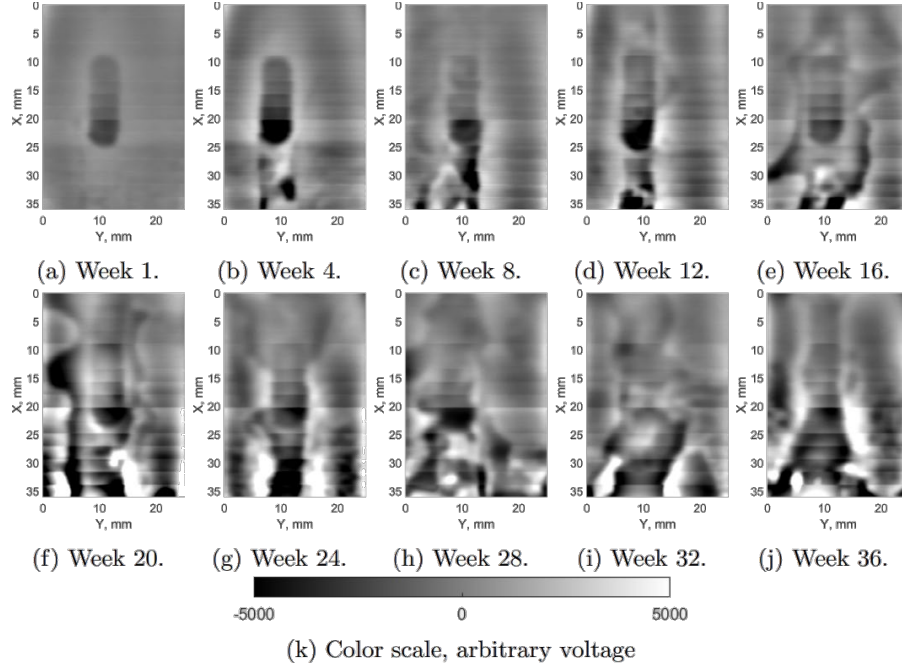


Figure 13. 2D post-processed data from real voltage component on corroded samples using the right column of coils.

The frequency used on the samples was selected at $100kHz$ as it was the resonance frequency of the probe to give the deepest readings under rust at around $920\mu m$. An input voltage at $v_{in} = 10V$ and pre-amp gain at $55dB$ was also used. Signals deeper than this range are expected to saturate, meaning more evaluation may take place from the surface to the end of the current density range. To calibrate gain per coil, the probe was placed $1mm$ away from the far-side, or with the sample flipped up-side-down, of week 1's sample. Magnifi's calibration tool was used to obtain gain settings, which only one calibration was used for all samples. A Shapeoko CNC gantry was used for scanning the sample, repurposed for raster scanning with sub-mm positional accuracy of the probe. Magnetic encoders were installed on the gantry to give high positional accuracy of the sensor array to the Ectane and Magnifi, which synchronizes both data position and values. The z-axis consists of a leadscrew to maintain a constant lift-off between of $1mm$ away from the surface. The sample was placed in a measured location on the gantry to enable consecutive scanning with minimal rotation of the sample or shifting. The setup is shown in Figure 14.

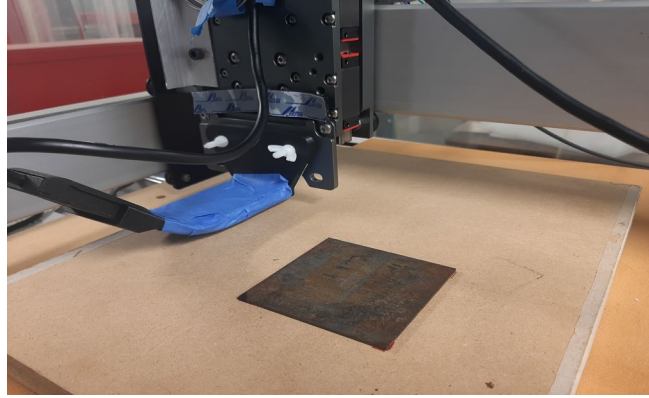


Figure 14. Eddy Current Array (ECA) NDE gantry setup.

Factors such as probe or sample tilt or overall sample thickness will vary lift-off. These effects may be cancelled out in post processing. A zig-zag raster scan was conducted creating an image with a resolution of $0.1 \times 0.1\text{mm}$. The scanning area is $36 \times 25\text{mm}$ to cover the sample. To obtain a 0.1 resolution along the shifting axis X , the gantry moves the probe in 0.1mm segments for 19 movements, between 0 and $2 - 0.1 = 1.9\text{mm}$, as the difference between 2 consecutive coils on a column is 2mm . If both columns were used for one image, then only 9 movements would be needed between 0 and $1 - 0.1 = 0.9\text{mm}$. For the scanning direction on axis Y , resolution is dependent on acquisition speed and gantry velocity. To prevent *null* data, the gantry moved approximately 15mm/s to provide consistent data, though higher speeds have been tested with minor data loss. Each scan took around 33 seconds to finish to obtain raw data, including buffer times to prevent exceptions between reading and writing the gantry, and excluding time placing the samples in and out of the scanning system.

Post processing operations are important for obtaining a clear image of the defects. The initial processing comes inside of Magnifi, which interpolates the raw eddy current data into a 2D voltage mapped image including real and imaginary components. In Matlab, the real data is detrended by creating a surface fit with polynomial order 5 along both axes and subtracting the results. The mean value of each 1D line is then subtracted, each along Y . After, a 10×2 median filter is used to remove speckle noise mostly in the scanning direction. The processing due to detrending placed the voltage into arbitrary units. For the 1D data, the results of the 16th coil, near the middle of the probe, was processed by averaging each shifting axis result. The middle of the probe gave the best indication of where the original defects occur. The data is detrended by subtracting the mean of the 1D result against itself. Finally, the data is zeroed by using the "no

defect" region on the right edge as a reference, and the results are shown in Figure 15.

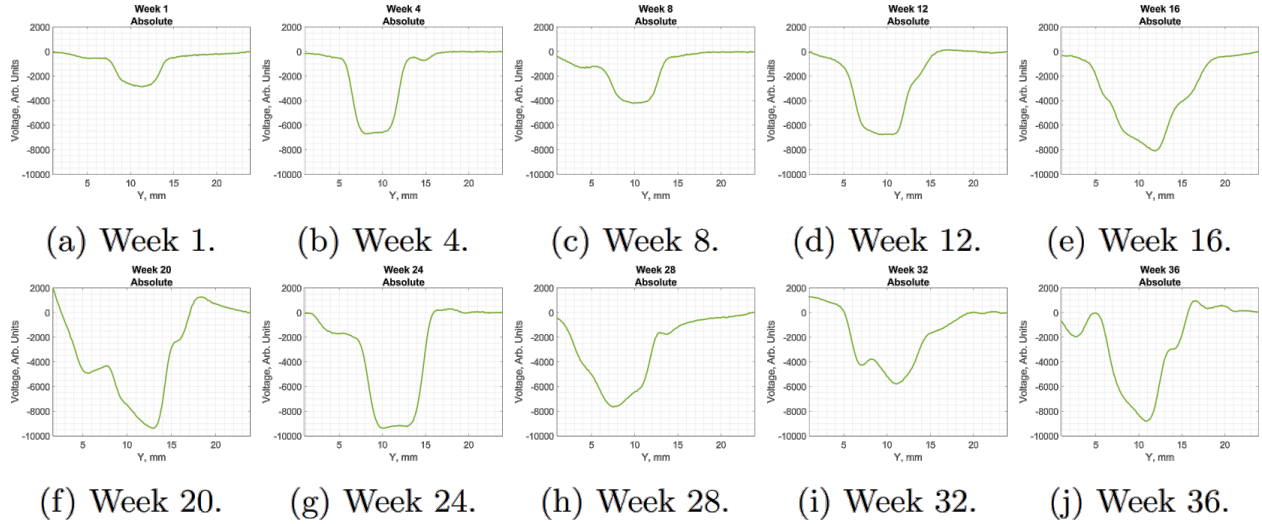


Figure 15. 1D post-processed data from real voltage component on corroded samples using coil 16.

From the 2D and 1D results, damage can be clearly seen. A pattern develops between week 1 and 36, showing the growth of corrosion with time. The original defect, which is seen clearly in week 1, deforms increasingly with time. Darker patterns on the 2D images indicate deeper damage, while brighter patterns imply corrosion defects above the surface of the sample. This effect has been discussed in other eddy current corrosion detecting literature [3]. Some saturation occurs in the later week samples due to the extent of corrosion, seen at the bottom of the 2D images between week 20 and 36. Some of this saturation may be due to differences in expected skin penetration, or even due to lift-off variation in calibration and scanning. The 1D results show drastic changes of depth pattern as corrosion increases with time. What is interesting is that depth does not consistently decrease with each passing week, only the increased deformations. There is also an increase in the defect's length along the Y axis.

There are some improvements that can be made in terms of the scanning procedure and post processing. This includes unbiasing each coil from the 2D results. There are "bars" seen in the 2D data sets, which represent the data collected from each coil. Differential mode scanning, where neighboring coils are subtracted, may help remove bias between coils. Precision tilt mechanisms may be placed on the sensor holder to calibrate tilt for these samples. An analysis can be done using gyroscopes and comparing rotational information between the sample and sensor may be useful in this regard. This would help the raw data from coil from being biased due to lift-

off variation. Later works will examine how to convert this voltage into depth, which is a well-known problem [4]. This would help give quantifiable information on how much damage is occurring rather than a more abstract "arbitrary voltage reading". Overall, the data collected shows exciting potential even with improvements for future works.

4.2 Ultrasonic NDE methods for interactive corrosion detection and characterization

The overall objective for Section 4.2 is to establish an expanded NDE framework for interactive anomalies by probabilistic characterization of defect profiles. The objective of using ultrasonic NDE methods is to develop numerical models and techniques for simulating guided waves (GWs) in pipeline geometries that include the plate wave equation to determine dispersion of GWs. Multi-scale and multi-physics modeling, we primarily look into present techniques for simulating guided waves in pipelines that include the plate wave equation to determine dispersion of guided waves. Meanwhile, modeling defect accurately is crucial in the simulation studies, since the NDE responses based on the modeled defects will be used to optimize the sensor frequency. We have used finite element modeling (FEM) to accurately model and mesh defect geometry to study the resulting ultrasonic NDE response. Using FEM will help not only optimize sensor parameters (e.g., frequency), but also study the physics behind the interaction of guided waves with complex interacting defects, and the generation and reception of guided waves in pipelines.

SHM and NDE of pipelines using ultrasonics requires a good understanding of defect signal vs. no-defect signal. While experiments can be carried out to understand this response, one should use a large set of data to effectively understand the differences. It would be efficient to develop an array of numerical models, which can simulate different materials and structural conditions to obtain their corresponding ultrasonic response for the complex anomaly scenario. This can further be used to develop the NDE and SHM protocols. In the sections to follow, we show the successful propagation of Ultrasonic Guided Waves (UGW) in a pipeline using a 2-D FEM based model, and also model corrosion pits and look at its respective ultrasonic NDE response. The idea behind this is to come up with a model and find the optimum parameters like frequency, excitation etc., that can be then utilized directly in models with realistic defect profiles that is to be developed/generated by the research group at UAkron. It also gives a clear idea between a defect and a no defect response that is desired before experimental studies are carried out.

In order to build the numerical model, the COMSOL ® Multiphysics 5.4 software has been

used. Even though the final goal is to build a 3-D model, it is important to first understand the underlying physics behind the propagation of UGW in a pipeline. The 3-D models become computationally very costly for this purpose. Also, debugging and validating the results in a 3-D model are also much harder. Therefore, a 2-D axisymmetric model can be used where the symmetric nature of a cylindrical pipeline can be exploited. This approximates the wave propagating in the axial direction, which is sensitive to the circumferentially oriented defects. For axially oriented defects we employ wave propagating in circumferential direction. For our preliminary studies using axial and circumferential guided waves, we have considered Steel AISI 4340, whose properties are enlisted in Table 1.

Whenever an ultrasonic guided wave is propagated through a medium, multiple wave modes are generated. For large pipes we can safely assume the wave characteristics to be same for axial and circumferential waves. This is because the wave is dispersive in nature, and the number of different types of wave modes depends on the frequency and the thickness of the sample. The two primary modes though are the symmetric mode (S0) and the antisymmetric mode (A0). These two modes are generated at relatively lower frequencies. At higher frequencies, there are multiple wave modes, which makes detection and isolation of particular wave modes very difficult, and thereby also making detection of defect signatures harder as they might get buried in the signatures of various different wave modes. Basically, dispersion of the wave causes multiple wave modes that can make data interpretation incredibly hard. Therefore, only the A0 and S0 wave modes are generated, and these signatures are observed in samples with and without defects. It is therefore very important we have information about the different wave modes and the frequency they are generated at for a particular material and thickness. This information can be deduced by looking at the dispersion curves for Steel. Figure 16 shows the dispersion curve for Steel AISI 4340.

Table 1. Properties of Steel AISI 4340

Density	7850 kg/m³
Young's Modulus	205 GPa
Poisson's ratio	0.28

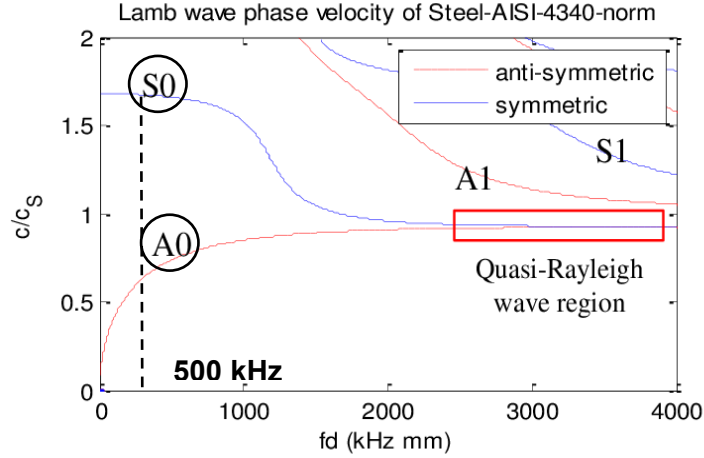


Figure 16. Dispersion curve for Steel AISI 4340

We have chosen 25 KHz as the operating frequency for a pipe with a wall thickness of 20 mm for which the modes generated are indicated in Figure 16. As desired, at this frequency and thickness, we will be able to generate only the A0 and S0 modes.

The first step is to understand the ultrasonic NDE response from a pipeline without any defects, as it is important to establish a reliable baseline where there are no defects. Figure 17 shows the sample geometry that is being considered. Since the goal of using ultrasonic guided waves is to perform long range ultrasonic testing, a 2 m long pipe was considered with a 20 mm wall thickness as mentioned before.

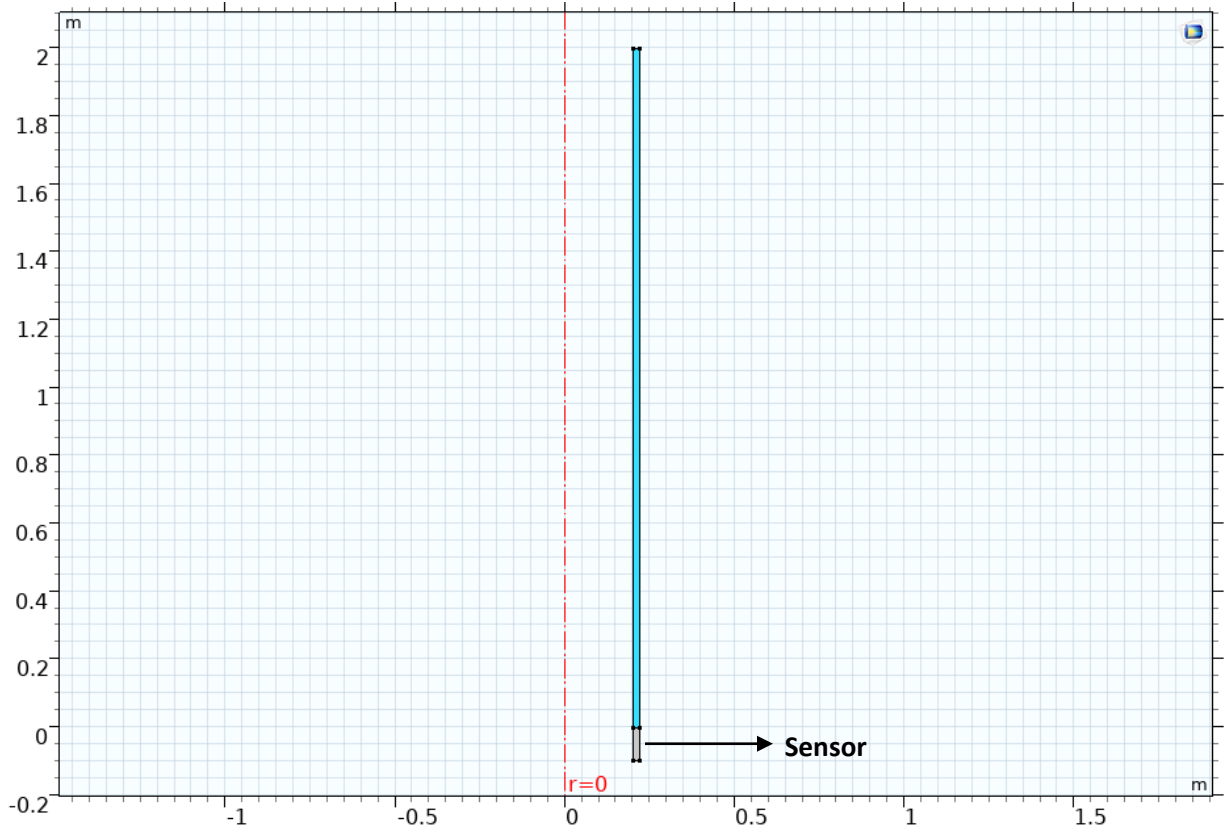


Figure 17. Sample geometry for 2-D axisymmetric model to simulate axial waves in pipes.

Figure 18 below shows the time domain representation of the excitation pulse used in this study. It is a typical ultrasonic tone burst signal modulated at 25 KHz with 10 cycles where it consists of a simple cosine signal modulated by a Hanning window. While Figure 19 shows the frequency domain representation of the burst signal, it is clearly observed that the peak lies at 25 kHz, which confirms the velocity of our excitation signal. Mathematically, it is represented by the equation below:

$$excitation = Sin(wt) * \left(1 - \frac{cos(wt)}{n}\right) \quad (4.2)$$

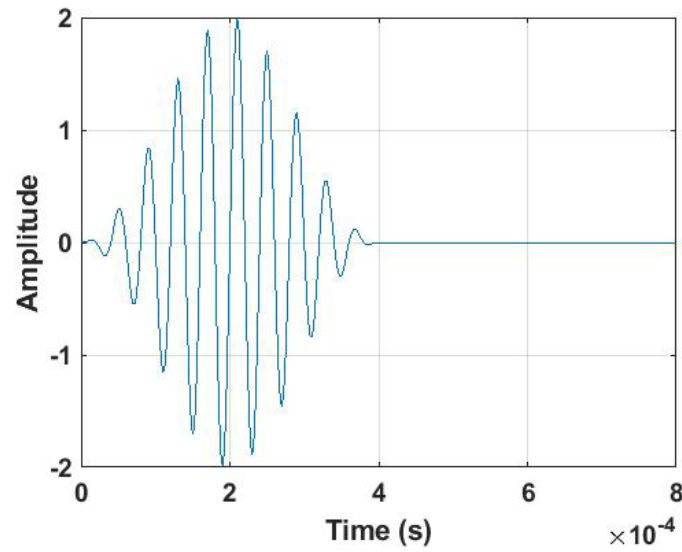


Figure 18. Time domain representation of the burst signal

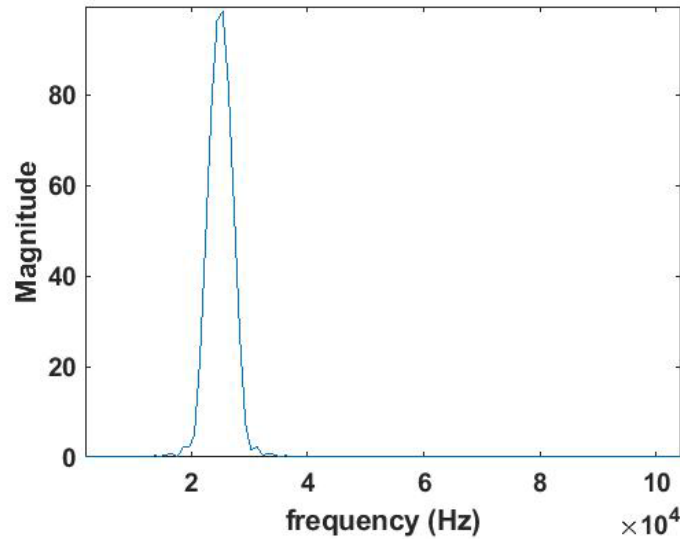


Figure 19. Frequency domain representation of the burst signal

The incident displacement is applied as shown in Figure 20 below. This way, the disturbance is applied across the whole circumference of the pipe. The quadrilateral (QUAD) elements available in COMSOL ® are used to mesh the whole domain. Also, a very fine mesh is used here where the minimum element size is 42 μm . A time dependent analysis using the direct linear solver MUMPS available in COMSOL ® is used to simulate the ultrasonic guided wave in the pipeline. The total time span for the simulation is 2000 μs while the time step is 2 μs . The degree of freedom in these

simulations is 2026.

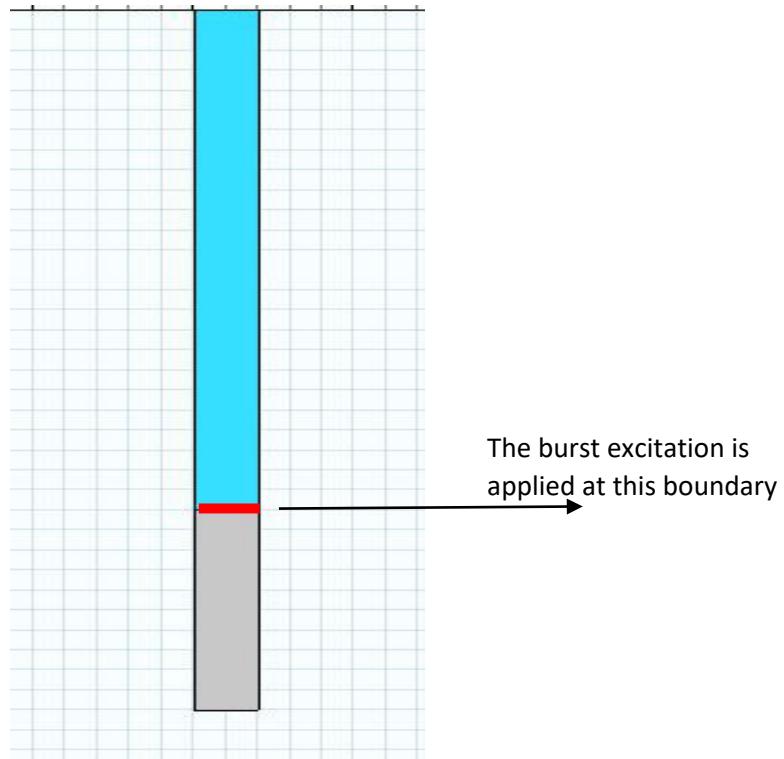


Figure 20. Zoomed simulation model showing the excitation by the application of the burst type signal on transducer boundary.

Now, looking at the velocity profiles of the ultrasonic guided wave, we can clearly distinguish and isolate the A0 and S0 modes. A point to note is that the transmitting and receiving points are the same, i.e., a pulse echo system was considered. Shown below in Figure 21 is the velocity profile or the A-scan where at the same point the disturbance was applied.

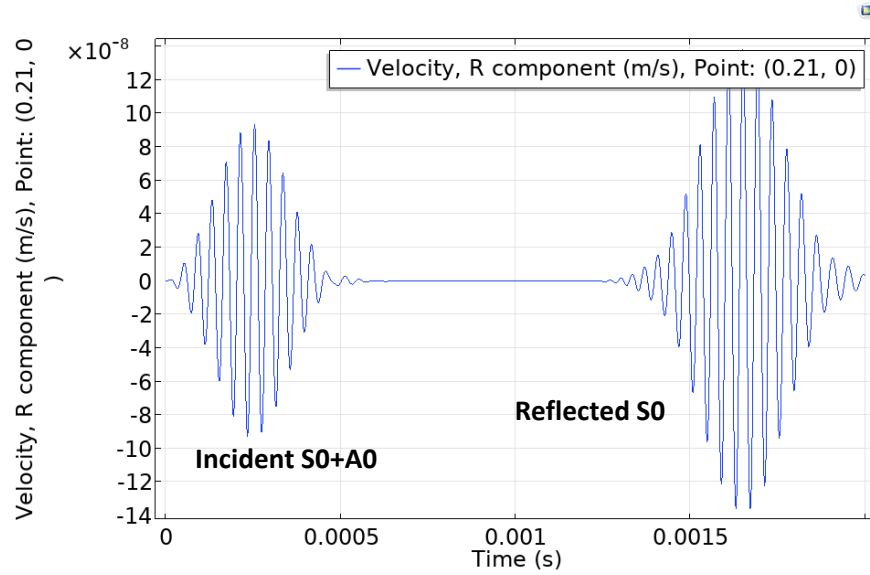


Figure 21. Velocity profile at (0,0) of the sample

From Figure 16, S0 mode has a higher velocity, therefore the first arriving reflected wave packet would be the S0 mode. The next arriving wave packet would be the A0 which is not shown. The difference in time in the arrivals of the incident S0+A0 and the reflected S0 mode is 1290 μ s. Figure 22 below, shows the resulting A-scan at a location 0.5 m from the starting of the pipe. In this case, the difference in arrivals is much lesser at 975 μ s. It is also clearly visible that the A0 mode arrives later i.e lower velocity and the S0 mode arrives faster i.e. higher velocity. This can easily be explained by looking at the dispersion curve in Figure 16.

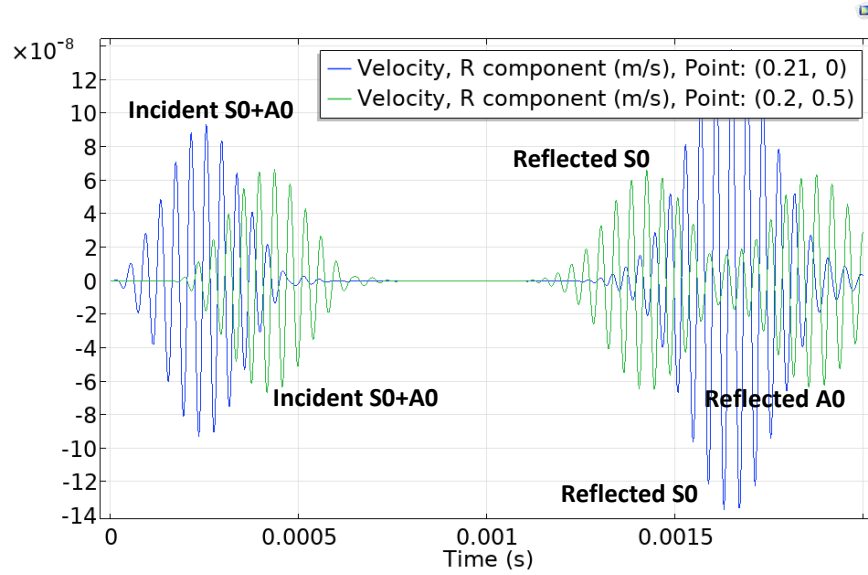


Figure 22. Velocity profile at (0,0) and (0,0.5) of the sample

A 3-D representation of the pipe in terms of the stress propagation is shown below in Figure 23. Since it is an axisymmetric simulation, the phi component is constant and the below figure shows the stress propagation at the end of the simulation i.e. 2000 μ s.

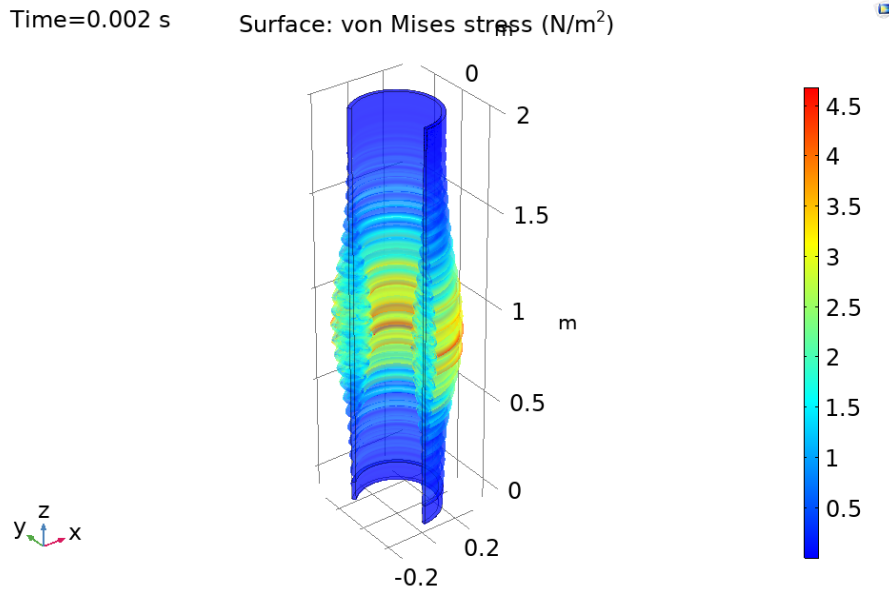


Figure 23. Von Misses Stress at time $t=0.002$ seconds

An ultrasonic guided wave has been successfully launched through the sample. The A0 and S0 wave modes are isolated easily and all the simulation parameters like the frequency,

excitation signal, mesh type and size, the type of solver have all been optimized. This model will serve as our baseline model for further analysis. The next step is to look at the ultrasonic NDE response in the presence of defects caused by corrosion, both internal and external. An understanding of the interaction of the propagated guided wave with defects is very essential for effective practical implementation. Pitting corrosion is one of the most dangerous forms of corrosion. Pitting corrosion is generally caused by environmental and material factors. For example, an abundance of chloride in the environment causes rapid pitting corrosion while inclusions in the material also aid in the process. Although there are many types of corrosion, only the modelling of pitting corrosion is undertaken this quarter. There are two main reasons for this, (1) the simplified damage models are easily generated for this type of corrosion and (2) it is the most common and dangerous type of corrosion occurring in metallic pipelines. Pitting is a localized phenomenon confined to a point or small area that takes the form of cavities. The combined effects of mechanical stress and pits severely affect the structural integrity of a pipeline. Pits can also very well act as sites for crack initiation. Generally, it is very difficult to accurately characterize the smaller sizes of corrosion pits.

Previous work has shown that corrosion pits can be easily modelled as hemispherical cavities on the surface of the pipelines. A similar approach has been followed in this work, where a hemispherical cavity is modelled on the surface of the pipeline. In the initial case, a single pit or cavity has been modelled at a distance 0.5 m from the origin. The length of the pit is 10 mm and its depth 1 mm. Figure 24 below shows the modelled pit on COMSOL ®.

Figure 25 gives the R velocity profile at the origin (0,0) and a 3-D representation in terms of the Von Misses stress. From Figure 25(a), we can clearly see some very small reflections arising in between the incident S_0+A_0 packet and reflected S_0 mode packet. This is elaborated in the following Figure 26 and Figure 27, where a clear defect signature can be obtained by taking the difference between the ultrasonic NDE response when there is a pit and the response when there is no defect.

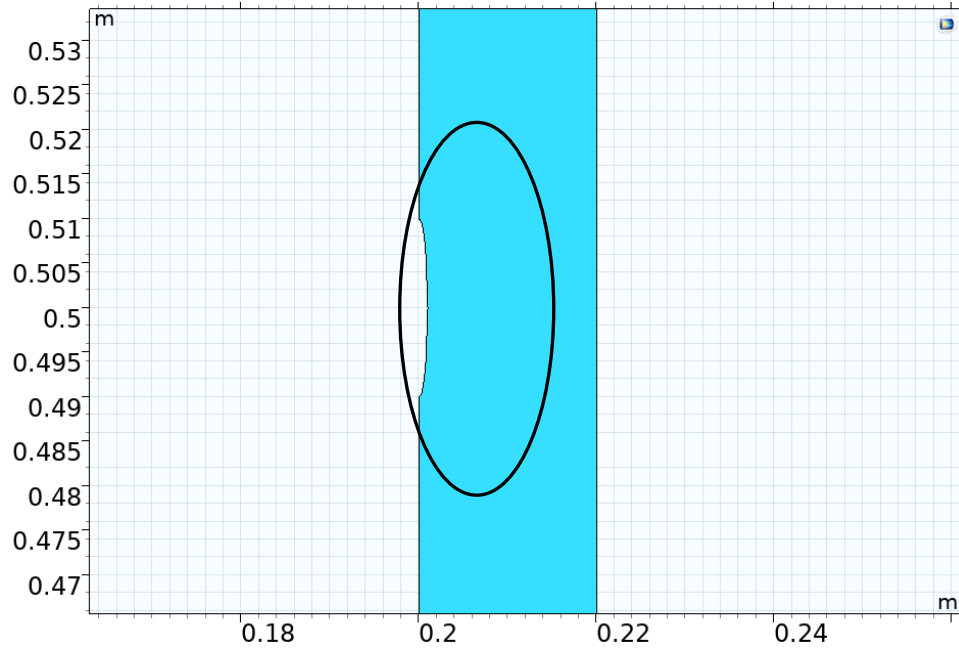


Figure 24. Cross section of the pipeline with a small pit

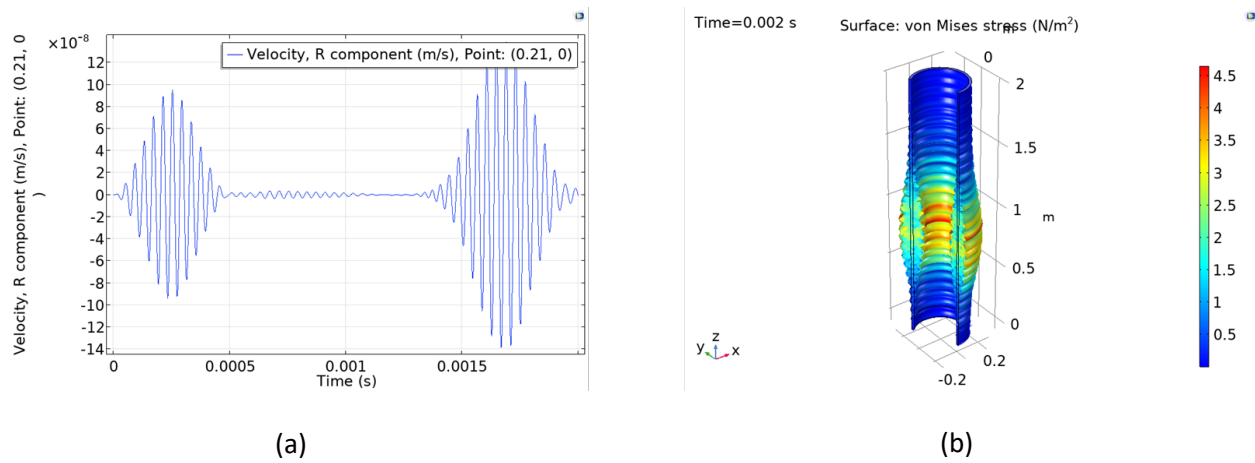


Figure 25. (a) A-scan at (0,0) and (b) Von Misses stress at time $t = 0.002s$

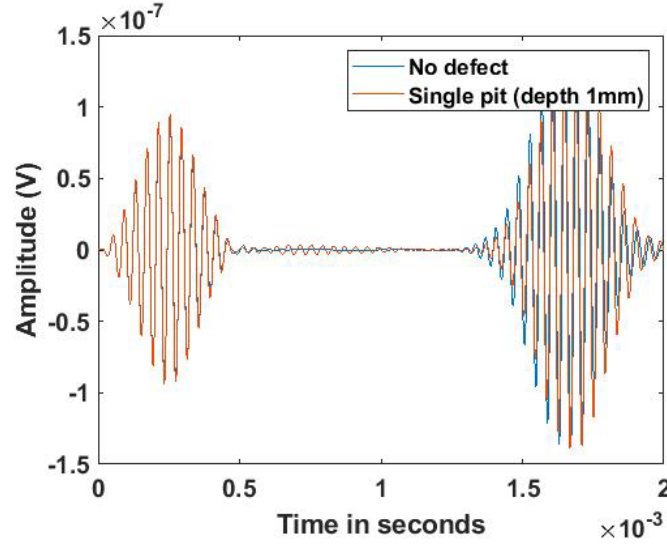


Figure 26. Comparison of A-scans with baseline

Clearly from Figure 26, it is observed the incident mode being free from reflection overlaps, while the reflected S0 mode shows clear difference in signature when there is a defect. The amplitude of the reflected S0 wave packet is clearly smaller when compared to the baseline model (i.e., no defect). This is attributed to the loss in the energy associated with reflected S0 mode from corrosion. Also, clear reflection though small pit is picked up in between the incident packet and the reflected S0 mode. Taking the difference between these two signals gives the defect signature arising directly from the defect. We have neglected the mesh noise to simplify the analysis. Figure 27 gives the defect signature.

The mechanics of corrosion and how it affects surfaces is a complex process. Hence, predicting the growth of pits requires extensive field and experimental study. The relationship for pit depth and time for a metal is loosely given as

$$d_{pit} = kT^{1/3} \quad (4.3)$$

where d_{pit} is the pit depth, T the exposure time and k is a constant based on the water and alloy composition. For this quarter, the objective was to conduct a preliminary study, and hence we manually simulated different models with different pit depths. Figure 28 shows the simulation results for a pit with length 10 mm and depth 3 mm.

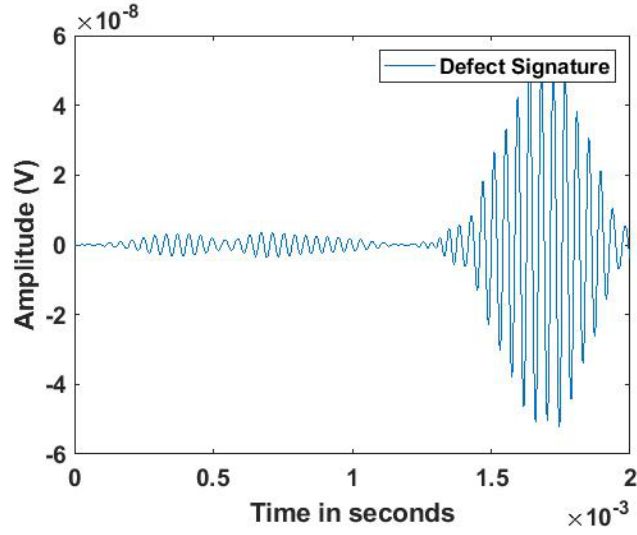


Figure 27. Defect Signature for pit with 1 mm depth

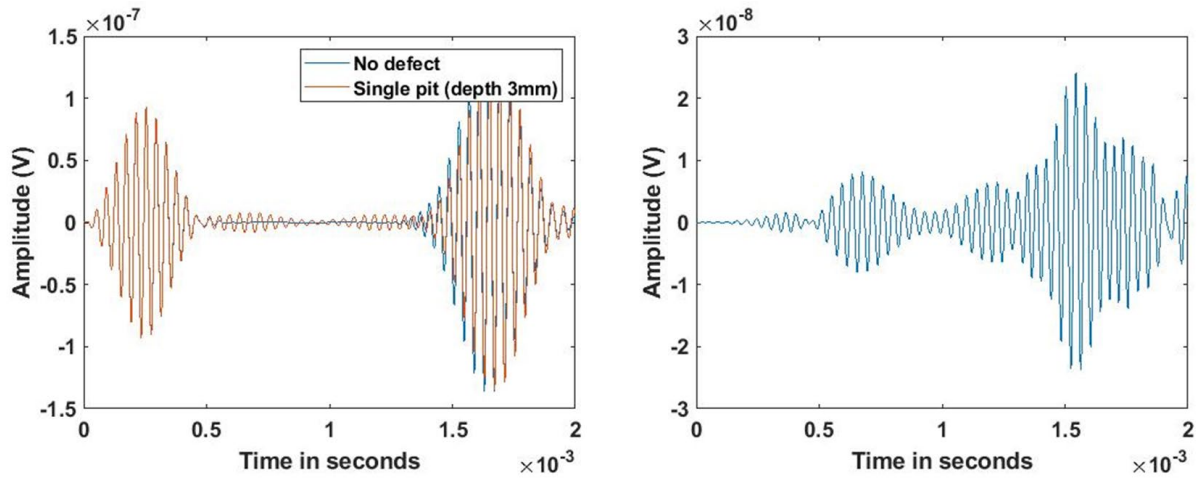


Figure 28. A-scans and defect signature for pit with 3 mm depth

As the pit depth increases, we see a clear increase in the amplitude of the reflections arriving between the two packets. The reflected S0 mode is also significantly smaller in amplitude as pit depth increases, thereby increasing the amplitude of the defect signature also shown in Figure 28. Figure 29 shows the signature for a pit depth of 5 mm with the length still being 10 mm. A similar trend is also noticed when the pit depth is 5 mm.

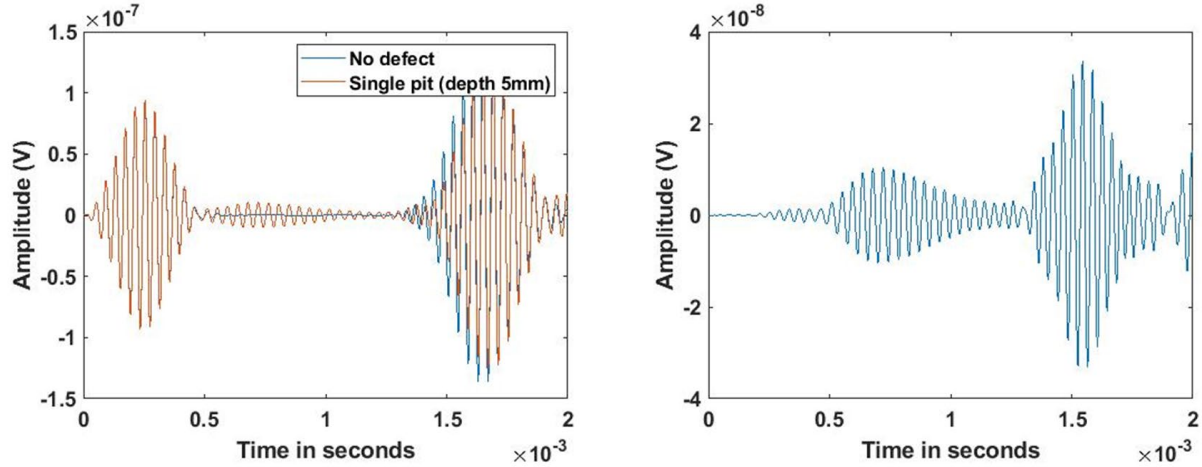


Figure 29. A-scans and defect signature for pit with 5 mm depth

In order to further quantify the different ultrasonic NDE responses obtained for different pit depths, Figure 30 below shows the defect signatures for pit depths 3 mm and 5 mm. The difference in signatures in terms of amplitude is very clear and could act as a viable feature during classification purposes.

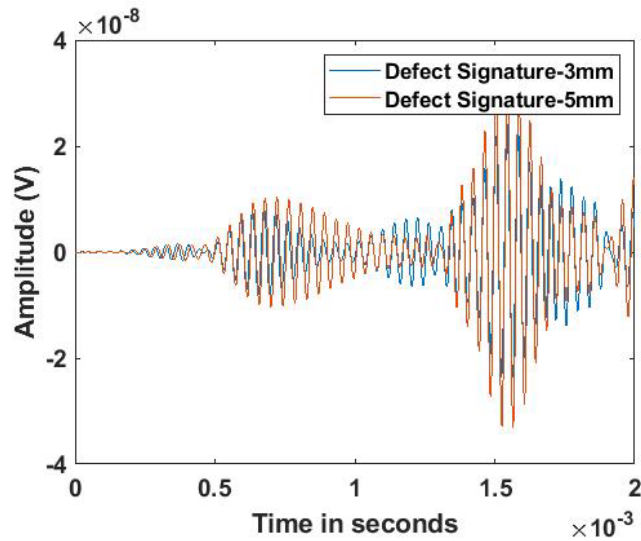


Figure 30. Defect signatures for pits with 3mm and 5 mm depths

Since the scope of this project is to look at novel NDE methods to understand interacting anomalies and the corresponding response, the next step was to look at multiple pits around the same area. To simplify this and completely understand the physics, two pits were modelled adjacent to each other. The model geometry is shown below in Figure 31. Both the pits are 10mm

in length and 3mm in depth.

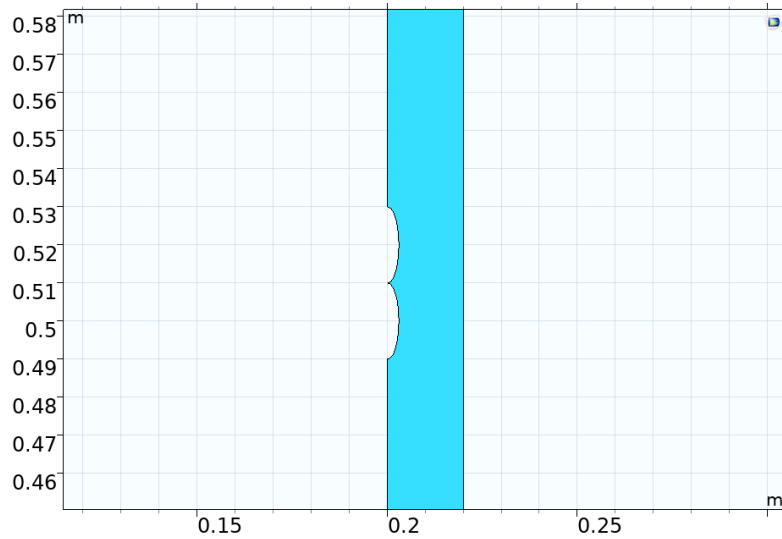


Figure 31. 2 pits modelled adjacent to each other on the surface

The A-scans and the defect signature is shown in Figure 32. In order to see a significant difference, the simulated A-scans here needs to be compared to the model with a single 3mm pit. Figure 33 compares the defect signatures for the two cases discussed. As expected, as the number of pits increases the amplitude of the defect signatures clearly increases, while the sheer shape of the wave itself is slightly different.

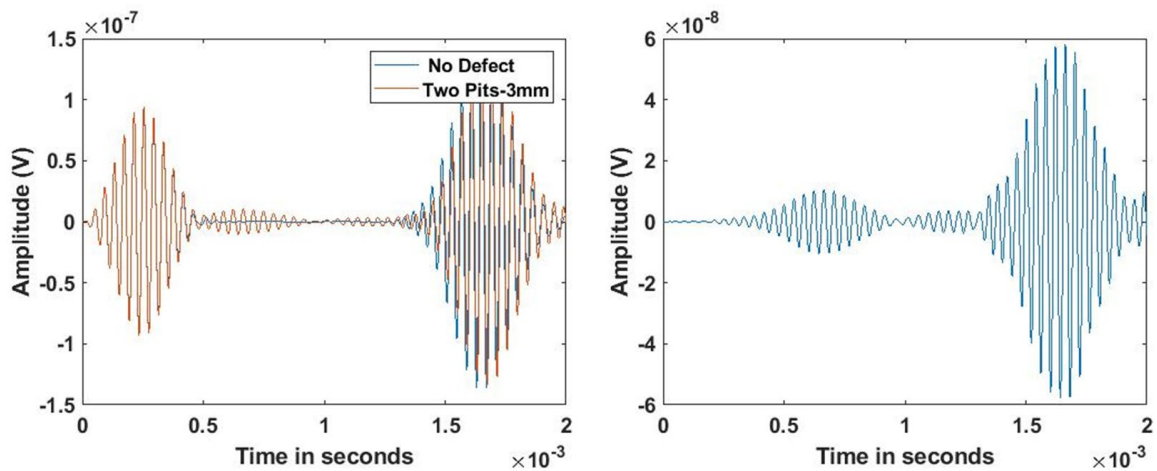


Figure 32. A-scans and defect Signature for two pits with 3 mm depth

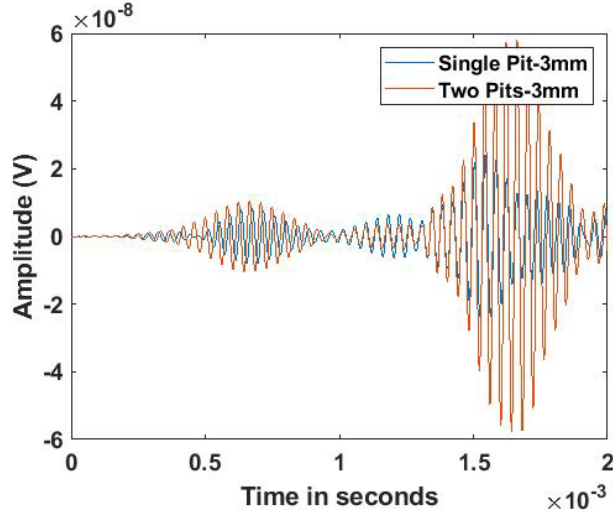


Figure 33. Defect signatures for single and double pit models.

We modelled a 2D asymmetrical model in COMSOL ® 5 Multiphysics software, where we modelled corrosion as hemispherical pits or cavities of certain depths and lengths. Figure 18 shows the burst excitation that is applied on the transducer boundary in our 2D asymmetrical model. The response for such corroded samples and clean samples were captured, and their simple difference gave rise to the pure defect signature arising from purely the corrosion pits. Also, since these pits/cracks normally exist in interactive colonies, the effect of number of pits has been studied. The material used was the normal Steel AISI 4340 usually found in many pipelines.

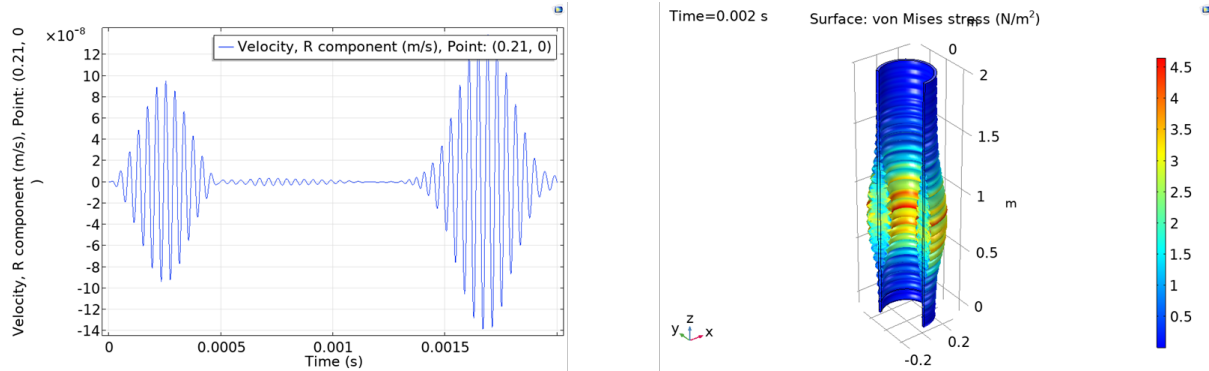


Figure 34. (a) A-scan at (0,0) and (b) von misses stress at time $t = 0.002s$

Figure 34 shows the velocity response of a pipe with an axial defect of 10mm in length and 1 mm in depth, while Figure 34(b) shows the propagation of Von Misses stress inside the 3D pipe. Figure 35 below gives the comparison of NDE responses between a healthy signal and a defect

signal. Clear difference in signals was observed for the pipe with a defect. Taking the difference between these two signals in Figure 35 gives the defect signature arising directly from the defect. We have neglected the mesh noise to simplify the analysis. Figure 36 shows the defect signature.

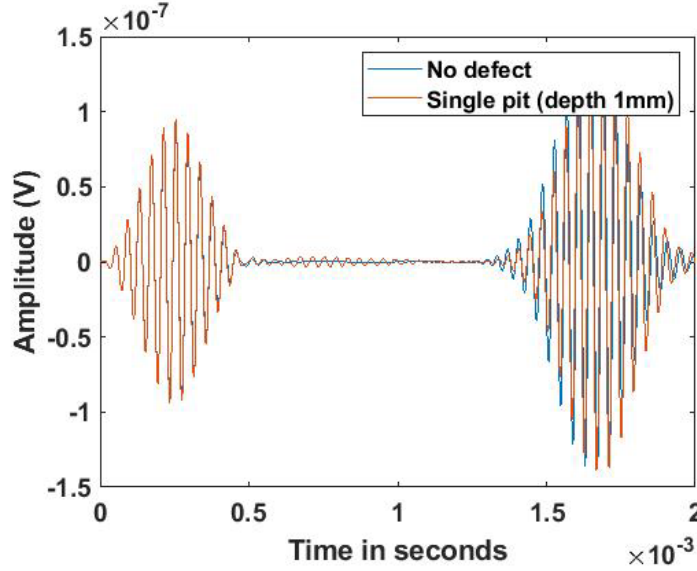


Figure 35. Comparison of A-scans with baseline

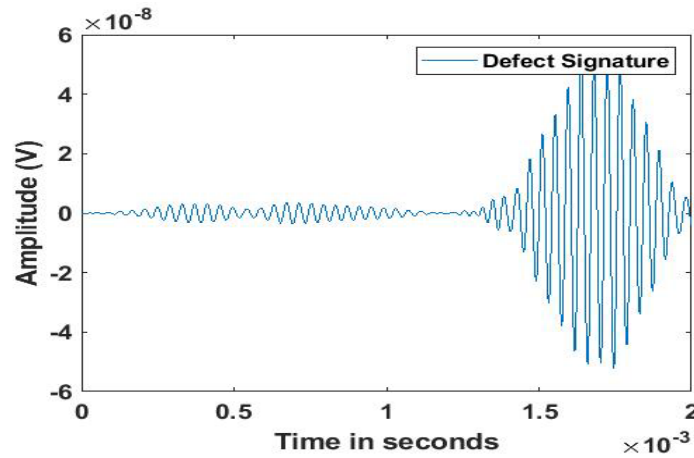


Figure 36. Defect Signature for pit with 1 mm depth

The mechanics of corrosion and how it affects surfaces is a complex process. Hence, predicting the growth of pits requires extensive field and experimental study. The relationship for pit depth and time for a metal is loosely given as,

$$d_{pit} = kT^{1/3} \quad (4.4)$$

where d_{pit} is the pit depth, T the exposure time and k some constant based on the water and alloy

composition. Figure 37 below shows the effect of pit depth on the defect signature. A clear increase in the amplitude of the defect signature was seen. This was expected, because as the pit depth increases, the reflections from the pits are much stronger, and since we used a pulse echo setup, the reflections are much stronger. A similar argument can be made if the number of pits increase, and this is clearly reflected in the defect signatures seen in Figure 38.

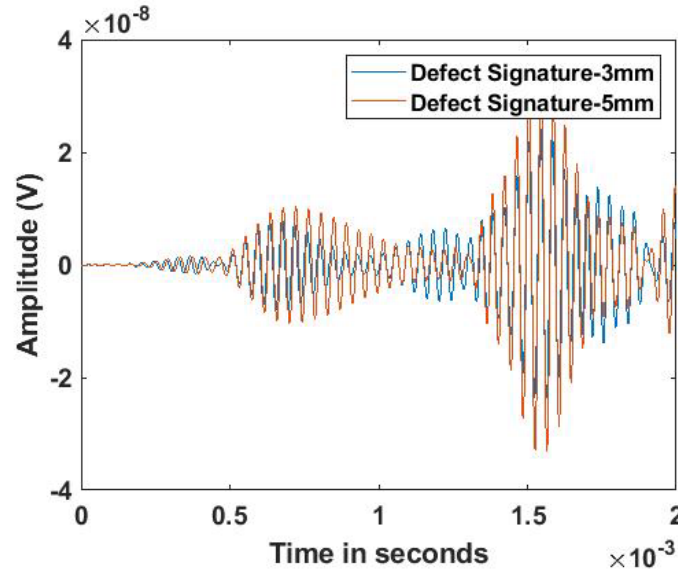


Figure 37. Defect signatures for pits with 3mm and 5 mm depths

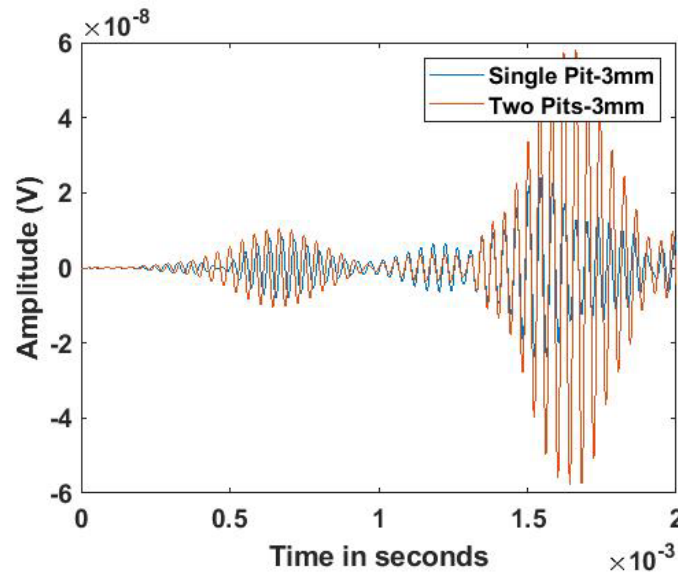
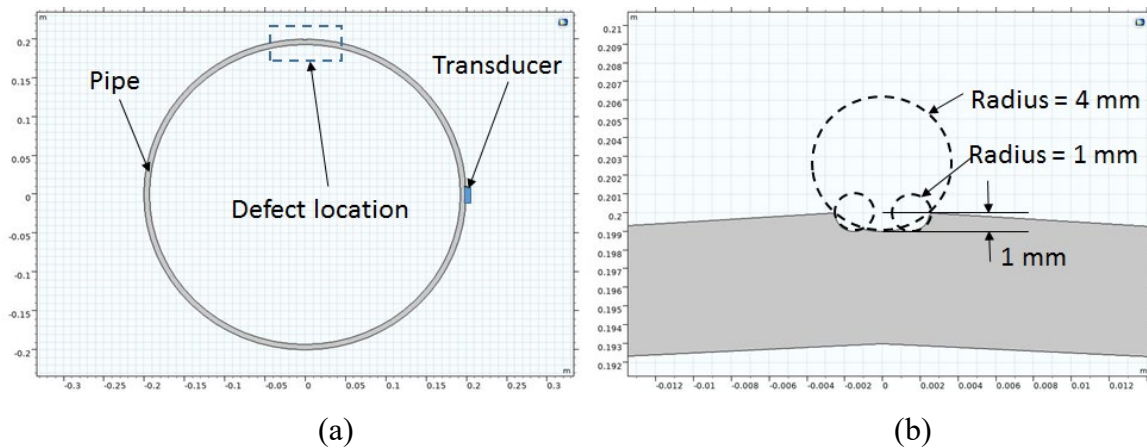


Figure 38. Defect signatures for single and double pit models.

4.2.1 Circumferential Guided Waves

Circumferential guided waves have the advantage of limited area to be covered depending on the circumference. Thereby, dispersiveness of the waves do not hold any limitation for the interrogation giving the operator freedom to choose any frequency. Circumferential guided waves are lamb waves that are launched using specific arrangement of transducers like the axially arranged phased array elements. Such waves are different than the one-dimensional waves in tubes. Commercial handheld scanners are available that utilize such linear array to scan the pipe length with its axial movement. We consider such a transducer as a point source in the study to study the circumferential guided wave interaction with corrosion. Consider a case of wave propagation along the circumferential direction at a frequency of 50 kHz in a 6 mm thick steel pipe with a diameter of 200 mm. The geometry and defect types are shown in Figure 39. The location of a piezoelectric wafer type transducer at 0° along the circumference and defect location is shown in Figure 39(a). Figure 39(b) shows the pit formed by the Boolean subtraction of three circles from the surface of the pipe. Figure 39(c) shows three pits with a central spacing of 6 mm. Figure 39(d) shows a simulated interactive defect formed by combination of corrosion pit and a crack. Colony of 3 pits with the location of crack in the central pit has been considered. The length of the crack is 2 mm deep. Detection of single and two pit colony has also been considered which has not been shown in the figure. The crack length in Figure 39(d) is changed to 1 and 3 mm to simulate the effect of crack depth on guided wave propagation.



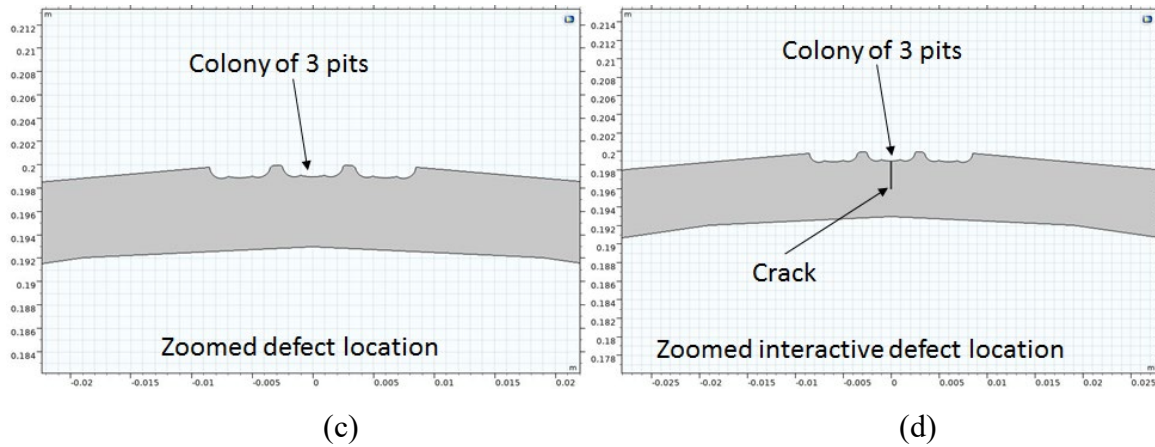


Figure 39. (a) Corrosion type defects located circumferentially around the pipe, (b) zoomed image showing dimensions of a single pit, (c) colony of three pits, and (d) interacting defects consisting of a colony of three pits with 2 mm deep crack in the central pit.

Excitation applied across the wafer type transducer produces S0 and A0 guided waves in the pipe that propagate around the circumference as shown in Figure 40. The pipe being defect free has a wave propagating in both directions from the wafer exciter. With a higher velocity the S0 wave propagates towards another end leaving behind the slower moving A0 wave mode. The top and bottom section of the pipe has similar wave propagation pattern due to symmetric geometry and transducer arrangement. The von Mises distribution is captured at 0.14 ms, which is enough to see the separation of the S0 and A0 wave modes. The wave modes would travel all the way around the circumference and reach the exciter when there are no defects or other structural features obstructing them. The signal received by the same wafer exciter in case of a healthy pipe is shown in Figure 41. The first packet is the incident packet appearing when the wave is launched by the exciter itself. Following packets are the S0 and A0 wave modes respectively returning to the exciter after propagating through the circumference. The signals obtained in the presence of corrosion pits have been superimposed. Clearly the signals vary due to the reflections of S0 wave mode from the corrosion. The reflected packet from the A0 wave mode is mixed with the returned S0 and A0 wave response and requires further processing to obtain it. The single pit produces enough change in the signal to be detected in the presence of a real environment with ambient noise. The reflected wave packet amplitude changes with its spread as the pits increase. This serves as a good indicator of damage severity which can be estimated by a cumulative damage index.

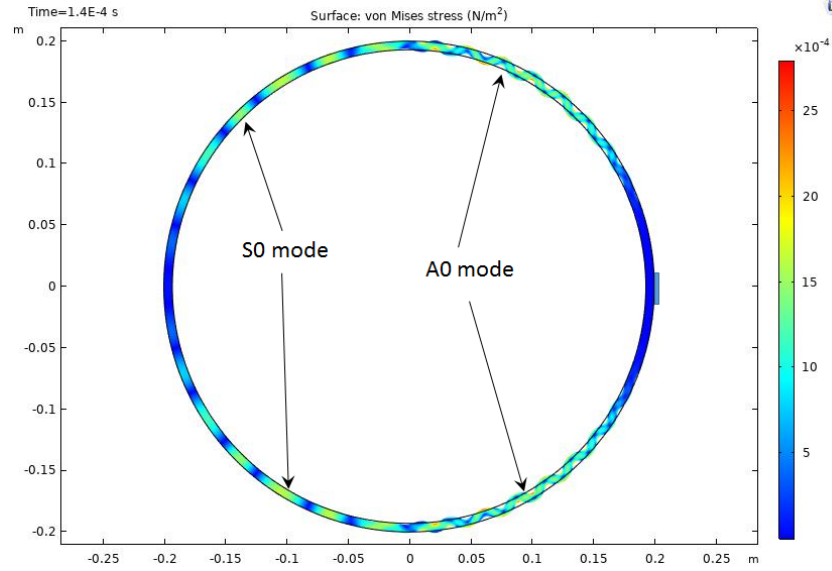


Figure 40. Circumferential guided wave around the pipe seen from the von Mises stress profile across the cross section.

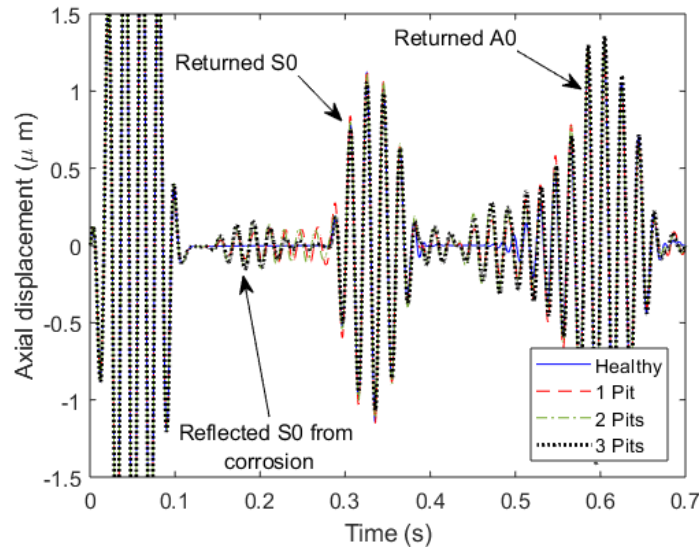


Figure 41. Axial displacement response at the piezoelectric wafer transducer location with healthy, one pit, two pit colony and three pit colony conditions.

Assuming that the three-pit colony is enough to introduce stress corrosion cracking, a crack is introduced at the middle pit with depth of 1 mm. The crack depth is further increased in steps of 1 mm to obtain another two cases of severity. Such interactive damage produces signals shown in Figure 42. The reflected S0 wave mode packet significantly increases in amplitude. The reason is attributed to the reduced cross section causing proportionate reflection of S0 wave energy. The

change in amplitude and waveform with frequency content can be further studied for damage classification including interactive features.

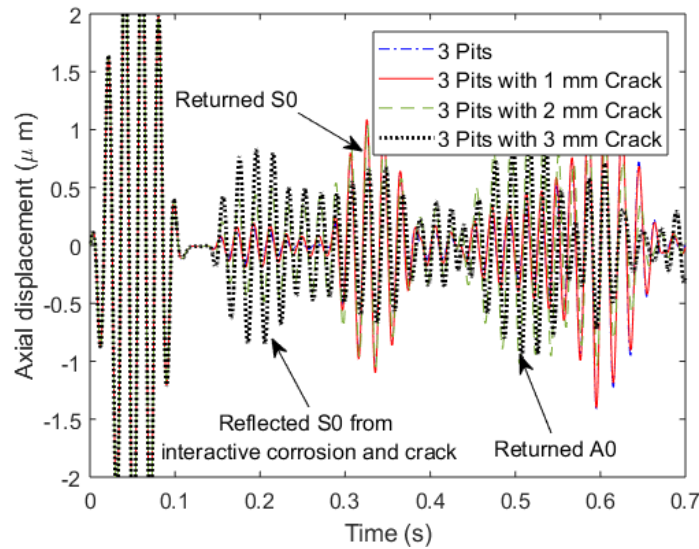


Figure 42. Axial displacement response at the piezoelectric wafer transducer location with three pit colony, interactive 3 pit colony with 1 mm crack, interactive 3 pit colony with 2 mm crack and interactive 3 pit colony with 3 mm crack conditions.

4.2.2 Automated Signal Classification

Simultaneously, the next step was to automate this process of feature selection and have a classification algorithm that can effectively predict the defect characteristics given a waveform. The advent of Machine Learning in signal processing, and especially in Nondestructive Evaluation has greatly helped this purpose. Neural Networks in brief are known as universal function approximators. But, for a complicated mapping, an exponential number of hidden units are required but such a large neural network may fail to train. Telgarsky investigated the importance of depth in neural networks. Deep neural networks encode a general belief that every function can be represented in terms of simpler functions and their combination can approximate the existing function. The underlying features can be extracted from the signal which has reduced a hectic and a time-taking feature engineering process. Each deep learning algorithm has its own pros and cons for wave response as features and is investigated by Rautela and Gopalakrishnan. Hence, choosing the right framework, architecture and the hyper parameters is a challenging task in itself. Deep learning techniques work by feed-forward propagation of input information to hidden layers to get some output. This output is not necessarily a true output (in a supervised learning setting). A back-

propagation algorithm flows information backward (which is generally a loss value described by a cost function) while using a gradient descent-based optimization algorithm. During the procedure of continuous forward and backward passes, the learning parameters (weights, W and biases, b) are tuned to a value that minimizes the cost function.

Current literature is bent towards the abovementioned optimization schemes but here, we have focused on using the Adam optimization. Adam is an adaptive learning rate optimization algorithm that's been designed specifically to train deep neural networks. Adam is a combination of RMSprop and Stochastic Gradient Descent (SGD) with momentum. It utilizes the squared gradients to adaptively scale the learning rate like RMSprop as well as the moving average of the gradient (instead of the gradient itself) like SGD with momentum. A neural network-based learning algorithm maps feature space to target space by minimizing the loss function using an optimization scheme (Adam optimizer here) over a virtual surface created by the dataset in n -dimensional vector space. A typical loss function is the mean-squared loss function (MSE). The formulation is presented in the equation below,

$$J(W, h) = \frac{1}{m} \sum_{m=1}^m L(y, \hat{y}) \quad (4.5)$$

A very important aspect while building such automated classification schemes is to collect a good 'distinguishable' dataset. By 'distinguishable', the dataset should contain signals that have features which vary for different conditions. Any neural network at the end of the day is a function approximator, and if there is no function to approximate i.e., when similar data is present, the networks fail. An important precursor to building accurate classification algorithms is to check the data for different features, and clean it if necessary. The features can range from simple features like temporal energy, peak amplitude, and time of flights to relatively more complicated features derived from Wavelet Transforms, Fourier Transforms and more.

4.2.3 Dataset Analysis

For this work, we only focus on the axial defect models for the dataset generation. We collected about 150 samples, which each collected individually were using the COMSOL models. The split up of the 150 samples is as shown below in Figure 43. Data collecting was a time-consuming process as each simulation took about 10 minutes, and 150 simulations adds up to about 25 hours of simulations. Hence our dataset is limited when it comes to the actual size. The simulation parameters are the same as reported in Section 4.2.1. The faulty signals were collected for range

of pit depths and lengths. Also, responses with different number of pits were collected.

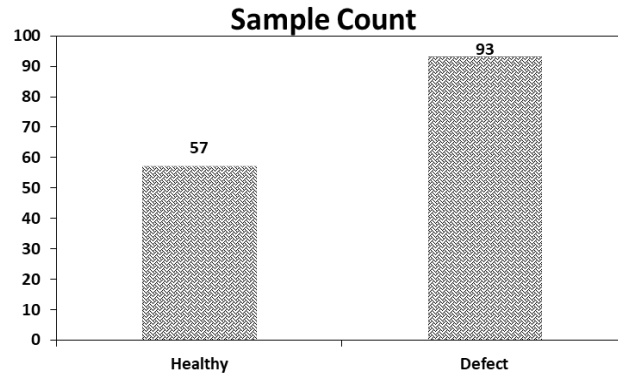


Figure 43. Split of the 150 samples collected for Classification

At first, we look at the simple statistical nature of the data. Simple features like the mean and variance of a vector is computed for all the 150 samples. Figure 44 shows the same. It is clearly seen that there is clear difference in statistics of the NDE responses of the healthy and defect signals.

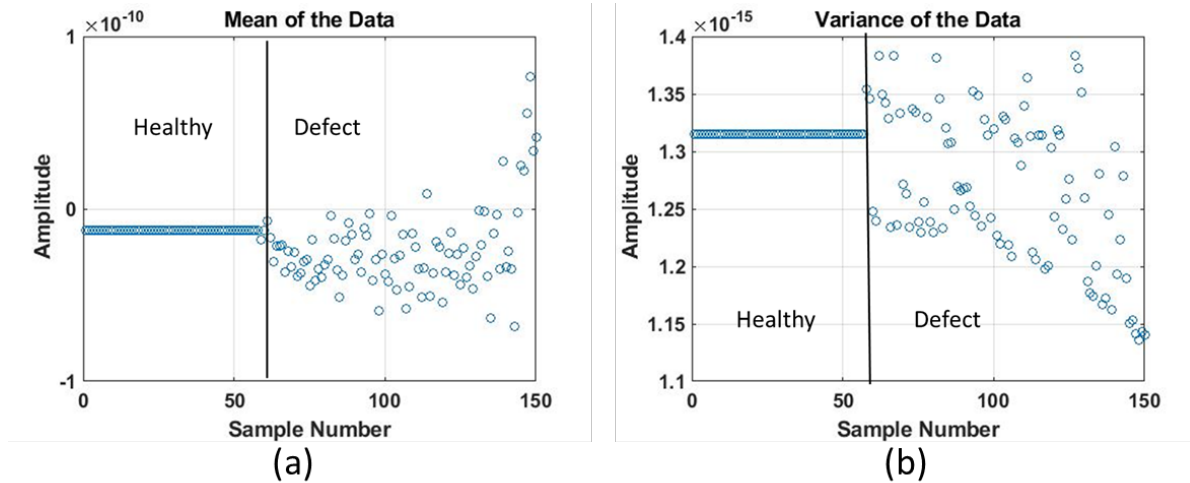


Figure 44. (a) Mean of each sample for the classification dataset and (b) variance of each sample in the dataset (150 samples)

Any NDE response is a time dependent data and also multiple frequencies at different times. Hence it is always important to analyze the temporal and spectral characteristics of such datasets. A simple way would be to look at the spectral and temporal energies of each signal. The

formulation for temporal energy is given below.

$$E = \frac{1}{2} \sum_{k=1}^n x_k(t)^2 \quad (4.6)$$

While the spectral energy formulation is as given below,

$$E = \frac{1}{2} \sum_{k=1}^n X_k(f)^2 \quad (4.7)$$

where $X_k(f)$ is the Fourier transform of $x_k(t)$ and is defined as,

$$X_k(f) = \sum_{k=1}^n x_k(t) e^{\frac{-j2\pi kn}{N}} \quad (4.8)$$

Figure 45 shows the temporal and spectral energy spread for each of the sample. Once again, considerable difference is seen for healthy and defect responses. Kindly note that all 57 responses of a healthy sample are expected to and has similar characteristics, both temporally and spectrally.

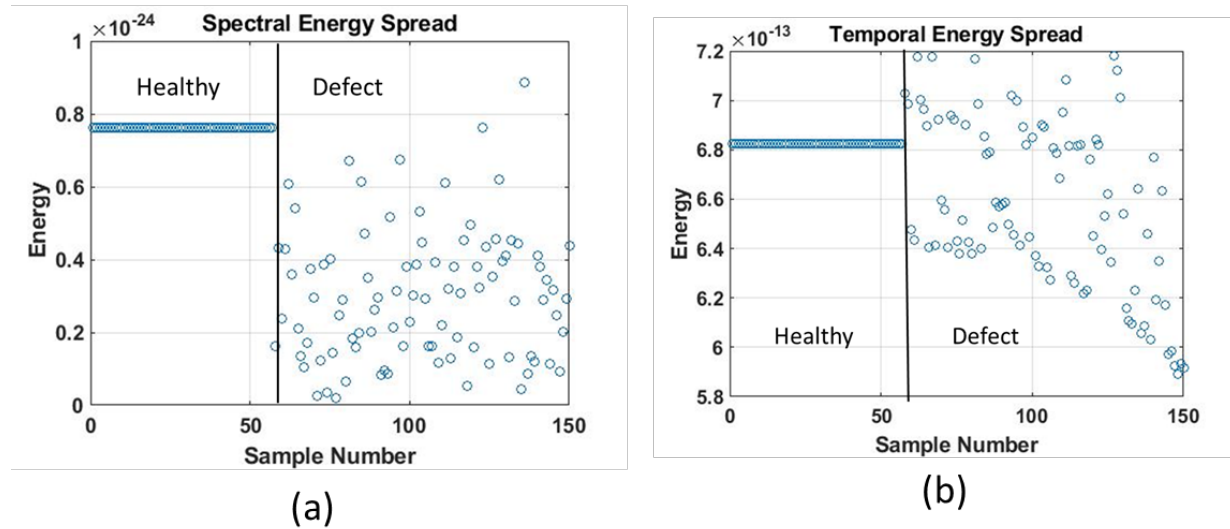


Figure 45. (a) Spectral Energy spread for the classification dataset and (b) Temporal Energy Spread in the dataset (150 samples)

The above analysis has given us a clear picture of the different features that can be

potentially used in the classification algorithm. The classification algorithm can be designed in two main ways. One method would be to feed in the raw A-scan itself, while the other would be to feed in the features as separate inputs post processing. The inputs can include the one's discussed above, or go beyond in terms of Wavelet coefficients, Wigner distributions. Feeding in only the features reduces the dimensionality of the problem, and henceforth makes sit computationally more efficient, while at the same time there is a risk if the features chosen don't really most accurately define the characteristics of a healthy or defect sample. This problem is avoided while feeding in the raw A-scan, but it makes it computationally more laborious. In our study, we have fed in the whole A-scan itself as the input, as it is not very clear from the study which feature influences the outcome most, and without that information, it would be very hard to choose the right set of features to train our network.

The first step is to be able to design a network, to simply classify healthy and defect signals in separate classes accurately. We have used a simple Multi-Layer Perceptron Network, whose architecture is described below in Figure 46. The network is trained on 145 samples of the dataset, and is tested 5 randomly chosen samples from the dataset. It consists of 4 dense fully connected layers, with dropout layers to avoid overfitting. By using dropout layers, we ensure the network works well not only to seen data, but also to unseen data. The network predicts a final value to be close to 0 or 1. It is considered a defect if it's close to 1, and healthy if it is close to 0.

Layer (type)	Output Shape	Param #
=====		
dense_105 (Dense)	(None, 1039)	1080560
batch_normalization_81 (Batch Normalization)	(None, 1039)	4156
dense_106 (Dense)	(None, 512)	532480
batch_normalization_82 (Batch Normalization)	(None, 512)	2048
dropout_53 (Dropout)	(None, 512)	0

dense_107 (Dense)	(None, 128)	65664
batch_normalization_83 (Batch Normalization)	(None, 128)	512
dropout_54 (Dropout)	(None, 128)	0
dense_108 (Dense)	(None, 1)	129
=====		
Total params: 1,685,549		
Trainable params: 1,682,191		
Non-trainable params: 3,358		

Figure 46. Architecture of the Multi-Layer Perceptron Network used.

The loss function used is the MSE function as described previously, while we have used an Adam optimizer. The activation function is Relu. Relu is typically used in neural networks to introduce nonlinearity in terms of the interaction of the inputs which is highly desired in practical problems. It is computed over 100 epochs, with a learning rate of 0.00001. Tuning the hyper parameters is a big aspect of building successful networks, and while there is no such right or wrong techniques to do so, it generally depends on the dataset and the architecture. We evaluate the performance of the network by the Mean Absolute Error (MAE) and the accuracy metric. One of the main takeaways during this was the significance of batch size while training. The batch size greatly influenced the performance of the network. The **batch size** is a hyper-parameter of gradient descent that controls the number of training samples to work through before the model's internal parameters are updated. Since our datasets was small, the chances the networks learn the same type of samples is a possibility. This will lead to over generalization of the problem, and the network won't predict well on unseen samples. Hence with a batch size of four, we were able to achieve very good performance as described below. The training loss curve and the training MAE plot are shown in Figure 47 and Figure 48, respectively.

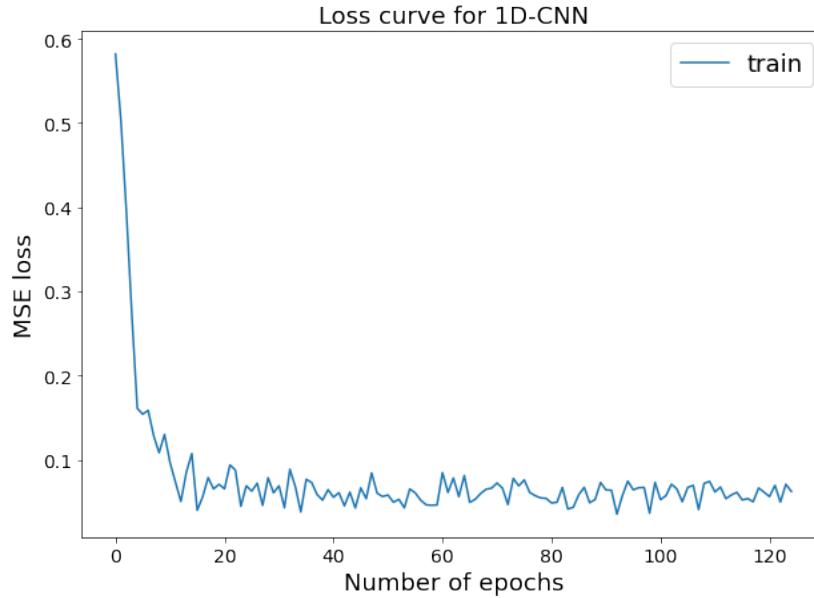


Figure 47. Training Loss for the MLP Network

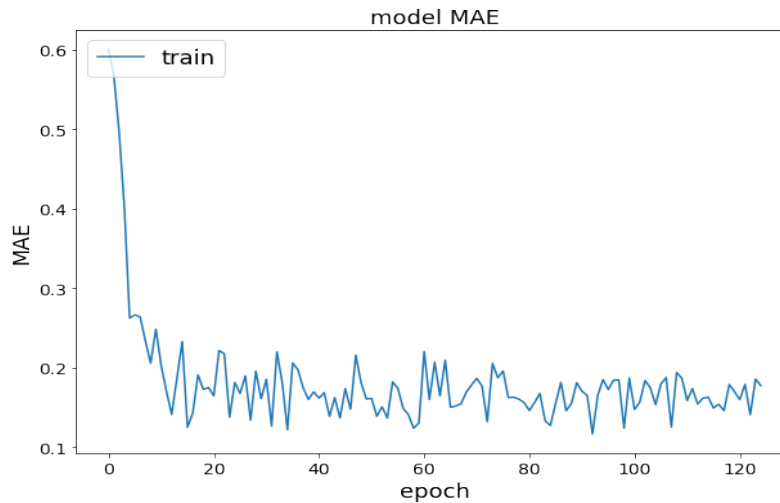


Figure 48. Training MAE for the MLP Network

The loss function converged pretty well during the training, while mean absolute error also converged well. The performance is obviously limited by the fact that our dataset is very small, and it is boosted by the depth of our network. We then used the network to predict the nature of the response on unseen examples, and it predicts with almost 96% accuracy. The accuracy across the training procedure is shown in Figure 49. The jagged nature of the plots in the previous figures (loss, MAE and accuracy) might be probably due to the size of the small dataset. Neural Networks

generally need a lot of data to learn and predict very well on seen and unseen examples, and in fields like NDE and SHM, generating or collecting big datasets is a challenge.

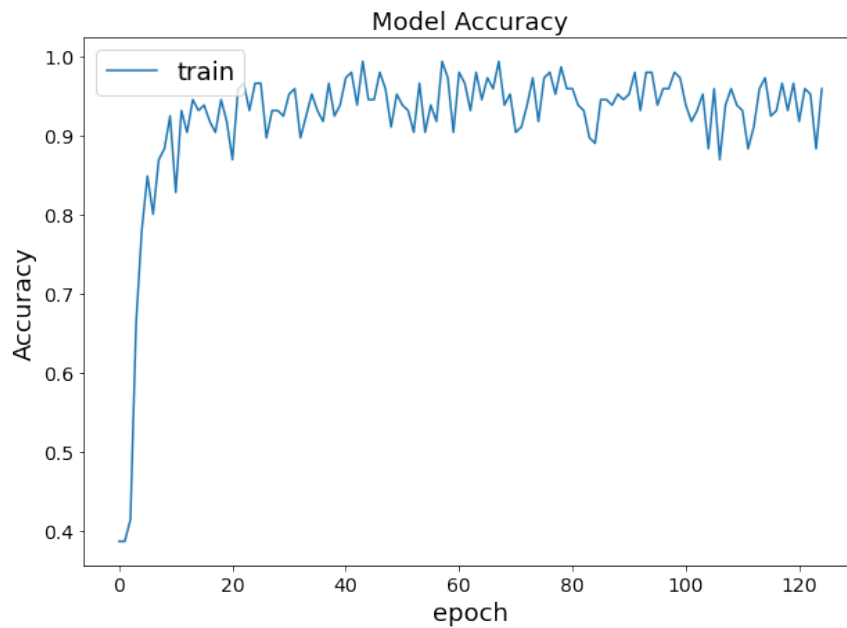


Figure 49. Training Accuracy for the MLP Network

The prediction results for the 5 samples are shown in Figure 50. The network correctly predicts four of the five samples to have defects in them, while it predicts correctly the only healthy sample. Please note that predicting a value close to 0 indicates a healthy sample, while something close to 1 indicates a faulty pipe.

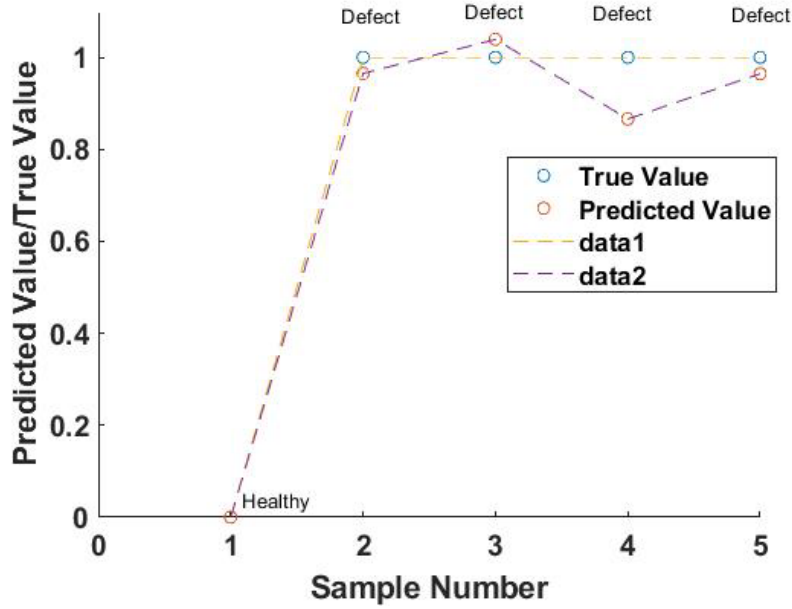


Figure 50. Prediction Results for the MLP Network

4.2.4 Conclusions & Discussion

We have been able to simulate circumferential ultrasonic guided waves inside a pipe with and without defects in order to optimize the necessary parameters. The interaction of circumferential guided waves in pipes with pits caused due to pitting corrosion has been captured. A comparison of signals with clear defect and a no-defect signal show capability to extract damage feature. The change in defect signatures with respect to pitting depth and number of pits is studied. Also, we were able to come up with an effective classification scheme to classify healthy and defect NDE responses with accuracy. Given any A-scan, the network can easily predict the health of the pipe. This opens up the prospect of using sensory data (either A-scans or images) to characterize corrosion pits completely. The next question is to develop an algorithm that can predict the characteristics of the defect in terms of its effective area, depth and location. For this, a much larger dataset needs to be collected and more rigorous signal processing schemes in terms of using Wavelet transforms among others.

The results obtained by simulating circumferential guided waves hold possibility of determining the damage parameters using a large dataset generated for different pit depths, pit lengths and number of pits, and for different transducer parameters like operation frequency and feature extraction techniques. The simulation results provide results to design a circumferential

transducer for experimental validation of these results are shown in Section 4.3. In terms of furthering a complete classification algorithm, the next step is to develop or improve the existing algorithm to predict defect characteristics like pit depths, length and location. Convolutional Neural Networks seem to work well with time dependent data and regression-based problems. The final classification scheme can characterize both axial and circumferential corrosion in terms of the depth, length, area, location and the number of pits. Looking at the bigger picture, other NDE methods like Electromagnetic NDE, Electromagnetic Acoustic Transducers are to be developed to not only characterize corrosion, but any pipeline related defects in general, and have a model that fuses data from different NDE modalities to predict most types of defects in a pipeline and characterize the defect in terms of its most basic characteristics.

4.3 Dataset Preparation for Machine Learning based Data Analysis

As discussed in Section 4.2, the dataset consists of 57 healthy responses and 93 defect responses. All the responses are collected from the axial guided wave model for corrosion pits. For simplicity, three different configurations have been considered. The first configuration is the healthy pipe, which is defect-free. The second configuration has one corrosion pit in the pipe, while the third configuration has a colony of two pits in the pipe. The length of the corrosion pits varies from 1 mm to 5 mm, while the depth of the pits varies from 5 mm to 20 mm. The operating frequency of the model is 25 KHz. In the last section, some basic features of the dataset were examined, which is shown in Figure 44 and Figure 45. The four features could be potentially used as feature vectors in various characterization algorithms. In this section, more complex features are added using some signal processing tools.

4.3.1 Cross Entropy Analysis

The Cross Sampling Entropy Method (CSamp-En) is mainly used to evaluate the degrees of asynchrony and dissimilarity of two time series in the same system. Please note that the responses collected in the dataset are velocity time histories. The CSamp-En method is based on the sampling Entropy method (SampEn) with a concept called Approximate Entropy that is basically a measure of the degree of irregularity or disorder in a measurement time series. When SampEn is lower, the sequence is more regular; the larger SampEn, the more irregular and complex the sequence becomes. SampEn is independent of the length of the data record and the algorithm uses the following parameters: threshold (r), sample length (m), and signal length (N). The length of the time series has no effect on the analysis results, and the results remain relatively consistent.

The analysis step of a CSamp-En algorithm is similar to that in the SampEn analysis method. The difference is that the object of SampEn analysis is mainly a single time series signal system, whereas the CSamp-En method analyzes two different time series signals to establish a template space for each of the two signals. This method basically gives a measure of how similar two time series signals are in terms of a similarity number. In the scope of our work, the similarity of different defect responses with respect to the healthy signal can be computed, which can be used to define a related Damage Index (DI). The procedure of CSamp-En is similar to that of SampEn and can be summarized as follows.

Let us define two time series signals: $\{X_i\} = \{x_1, x_2, \dots, x_i, \dots, x_N\}$ and $\{Y_j\} = \{y_1, y_2, \dots, y_j, \dots, y_N\}$. Both the time series signals are of the same length N . The two signals are

then divided into templates of size m :

$$u_m(i) = \{x_i, x_{i+1}, \dots, x_{i+m-1}\} \quad 1 \leq i \leq N - m + 1 \quad (4.9)$$

$$v_m(j) = \{y_j, y_{j+1}, \dots, y_{j+m-1}\} \quad 1 \leq j \leq N - m + 1 \quad (4.10)$$

A similarity number between $u_m(i)$ and $v_m(j)$ is defined as $n_i^m(r)$, and can be expressed as:

$$n_i^m(r) = \sum_{j=1}^{N-m} d[u_m(i), v_m(j)] \quad (4.11)$$

where, the maximum distance $d[u_m(i), v_m(j)]$ between the two template spaces $u_m(i)$ and $v_m(j)$ is defined as:

$$d[u_m(i), v_m(j)] = \max\{|x(i+k) - y(j+k)|\} \quad 0 \leq k \leq m-1 \quad (4.12)$$

$$d[u_m(i), v_m(j)] \leq r \quad 1 \leq j \leq N-m \quad (4.13)$$

When the distance between the two samples is smaller than the threshold, r , the two samples are considered similar; conversely, when the distance between the two samples exceeds r , the two samples are considered dissimilar. The threshold r can be chosen manually by the user. Through the use of different templates for similarity comparison and the calculation of the number of templates that exhibit the conditions of similarity, the number of similar samples in the i^{th} template to those in the entire template space can be obtained. The similarity probability of the i^{th} template can be calculated as:

$$U_i^m(r)(v||u) = \frac{n_i^m(r)}{(N-m)} \quad (4.14)$$

The average probability of similarity for template m can then be obtained as:

$$U^m(r)(v||u) = \frac{1}{N-m} \sum_{i=1}^{N-m} U_i^m(r)(v||u) \quad (4.15)$$

The degree of dissimilarity resulting from the division of the two time-series by m points represents the degree of synchronization between the two template spaces. Finally, the sample space is composed of the sample of length $m+1$, and the average similarity probability is calculated. The formula for calculating CSamp-En is expressed as:

$$CS_E(m, r, N) = -\ln \left\{ \frac{U^{m+1}(r)(v||u)}{U^m(r)(v||u)} \right\} \quad (4.16)$$

In this investigation, the changes of both final CSamp-En and the similarity number $n_i^m(r)$ are studied for different corrosion configurations and different corrosion pit depth. The parameters used in this study are listed in Table 2.

Table 2. Parameters for CSamp-En	
Signal length, N	1038
Template length, m	4
Threshold (based on healthy response), r	0

A DI based on the mean of the similarity numbers for different configurations is defined. The variation in the DI for different configurations is plotted in Figure 51, while the DI for different defect depths is plotted in Figure 52. From Figure 51, it is observed that there is a direct relationship between the number of pits and DI. As the number of pits increases in the colony, the absolute DI value increases, as shown in Figure 52. For defect depth, the DI value does increase for a defect depth of 5mm and 10 mm, but the DI values for defect depths of 15mm and 20mm are similar to that of 10mm. The increasing DI values basically indicate that the responses of a particular group are more dissimilar to the healthy response. This can be directly related to a conclusion that the samples corresponding to the particular group are relatively more damaged. For example, the samples with 5 mm defect depths have a smaller DI value compared to samples with 10 mm defect depths. This basically suggests that samples with 10 mm defect depths are more damaged, which is truth. A similar study is also conducted for changing the defect length (i.e., corrosion pit length).

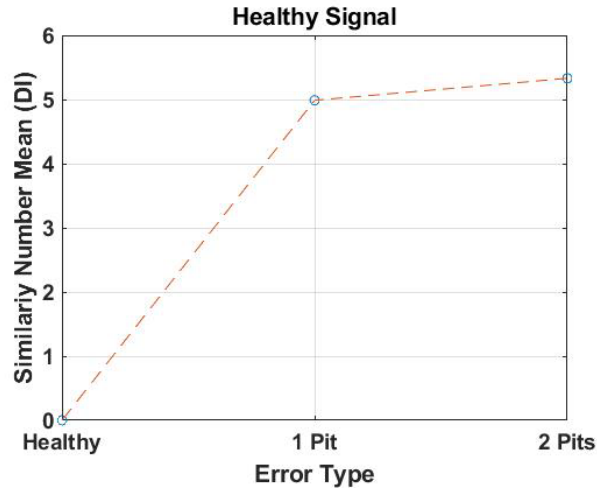


Figure 51. Normalized similarity number $n_i^m(r)$ vs number of pits

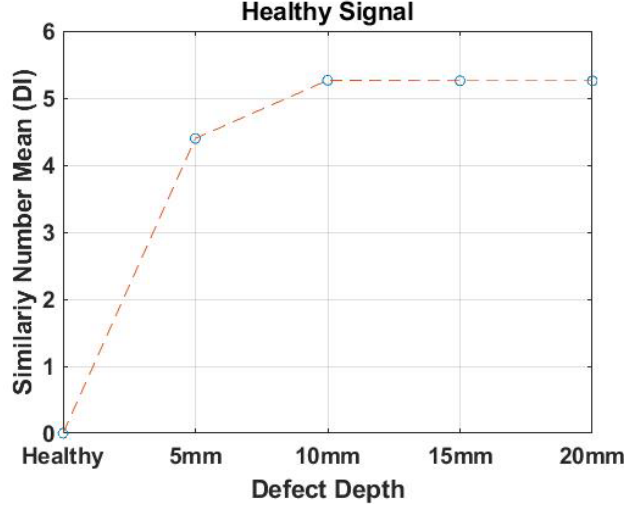


Figure 52. Normalized similarity number $n_i^m(r)$ vs defect depth

When using the CSamp-En analysis, we were able to extract a DI that shows a direct relationship to the number of pits. However, the results are not as promising for different defect depths. Nevertheless, overall the algorithm presented above provides a valuable feature engineering tool.

4.3.2 Hilbert Transform Analysis

Hilbert Transform (HT) is one of the most commonly used signal processing tools to study time signals. Implementing a HT enables us to create an analytic signal based on some original real-valued signal. HT gives the instantaneous amplitude of a signal, and it can be used to find the envelope of harmonic signals, which exactly is the nature of the response in our dataset. HT is mathematically described below for a signal $u(t)$:

$$H(u(t)) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u(\tau)}{t - \tau} d\tau \quad (4.17)$$

Similar to how MSCE is studied as a potential DI, HTs of three different configurations (i.e., healthy, single pit and colony of pits) are computed. The results are shown in Figure 53. The two peaks (i.e., Peak I and Peak II) shown in Figure 53(a) are basically due to the two reflections seen in the raw signal. The two reflections are the S0 wavemode, and a combination of A0 and reflected S0 wavemodes. As shown in Figure 53, one can observe that though the S0 peak (i.e., Peak I) has little difference for different configurations, the second peak (i.e., Peak II) shifts slightly to the right with increasing number of pits in the colony. The increase is in the order of 0.1-0.2 ms (as shown in Figure 53 (b)), which is significant considering the scale which we work

with.

HTs of the signals from different defect depths are also computed and the results are shown in Figure 54. Similar to Figure 53, there is no distinct difference among the HT responses at Peak I, but the second peak certainly shows a difference between the responses from a damaged pipeline and a healthy response. Therefore, Hilbert Transform is another valuable tool to compute a feature that distinguishes different configurations and defect depths. A DI based on the HT can be easily defined to quantify the corrosion defects, which is part of the work planned for the future.

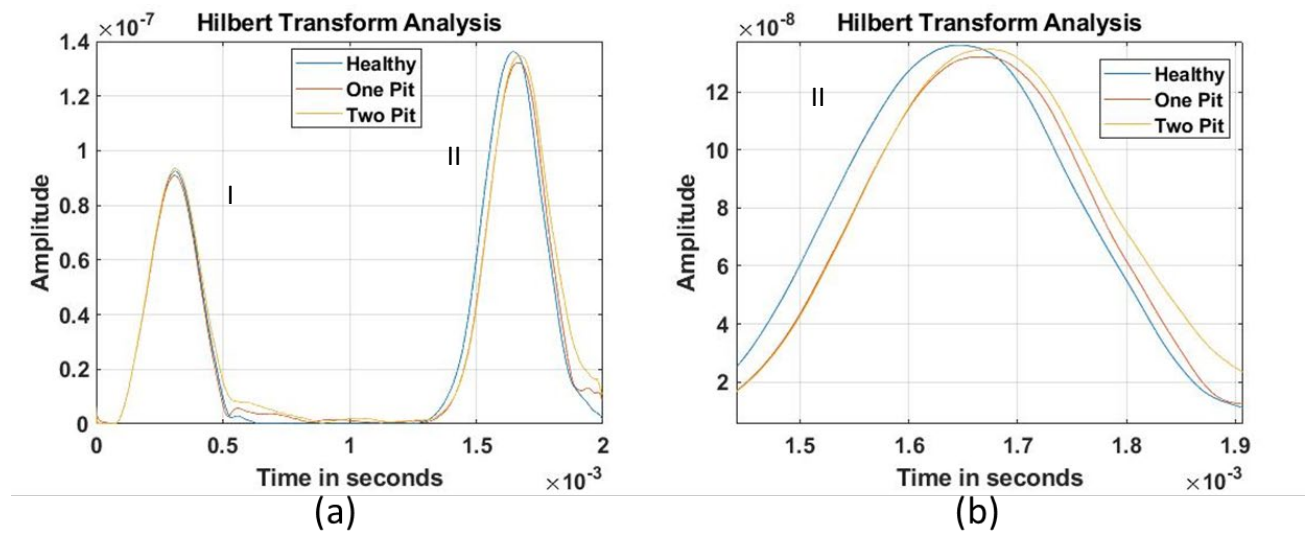


Figure 53. (a) HT responses for increasing number of pits and (b) a zoomed-in version of Peak II

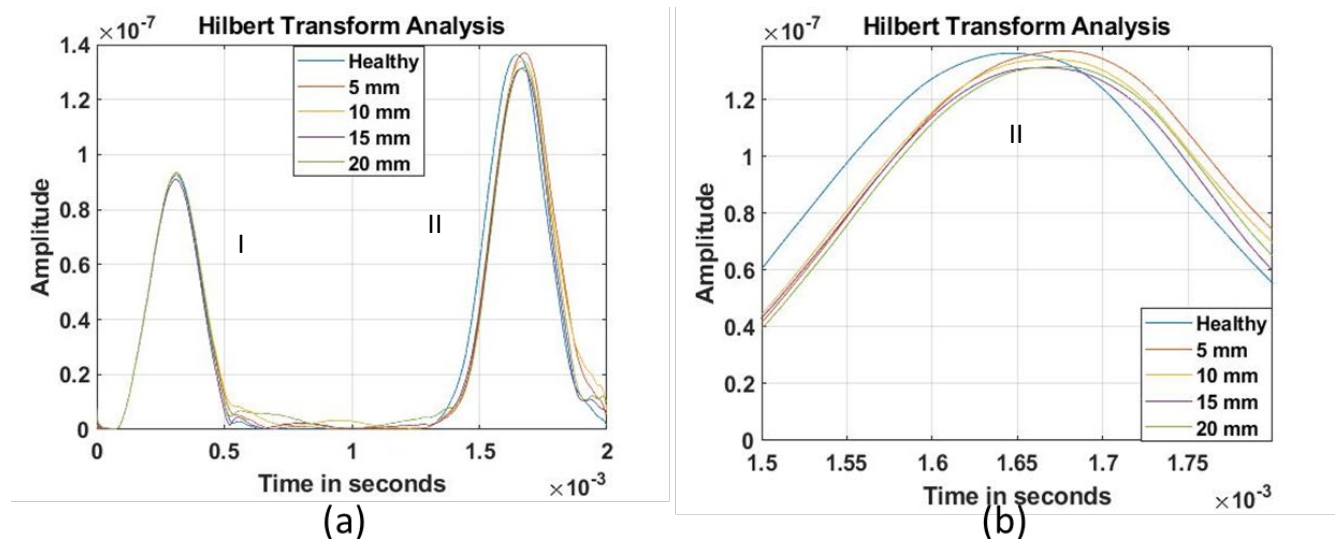


Figure 54. (a) HT responses for increasing defect depth and (b) a zoomed-in version of Peak II

4.3.3 Multi-Layer Perception (MLP) Classifier Network

In this section, we also tried to improve the robustness of the MLP based classifier network presented in the last section. For a quick recap, Figure 55 shows the architecture of the network that was used. Figure 56 shows the results when the trained model was used to predict on the unseen responses. For the dataset we defined above with well-labeled responses, the classification accuracy is defined as the fraction of correct predictions over the total number of predictions. The accuracy of this particular network was about 95%, which means that the MLP network performs really well.

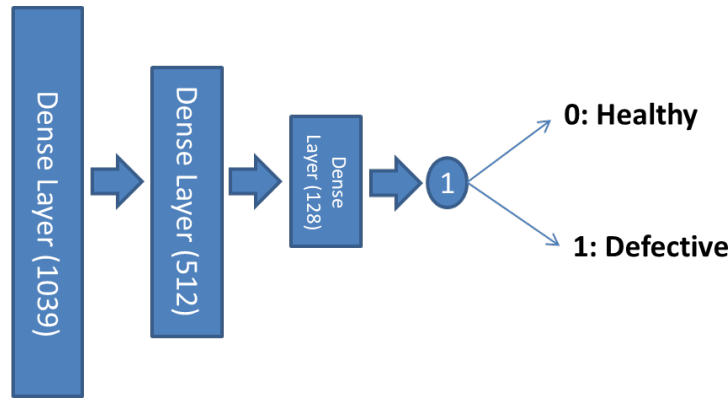


Figure 55. Multi-Layer Perceptron network for defect response classification

In order to simulate a real-world experiment environment, additive white Gaussian noise (AWGN) of different Signal to Noise Ratio (SNR) levels are added to the training dataset. The model is then trained, and predicted on the unseen examples. For this analysis, three different SNR levels of 5, 10 and 20 are considered. The signal with a SNR level of 5 has the highest noise content, while the signal with a SNR level of 20 has the least noise content. Figure 57 shows a signal without noise, and a signal with an artificially added noise of SNR 5. As shown in Figure 57 a signal with SNR 5 is highly distorted with noise, and training on such data is tricky and challenging. However, training with such high-level noises can make a model more robust and more generic in nature. As expected, the performance of the MLP drops as noise is added. Figure 58 shows the performance of the model when it is trained on a dataset with a SNR level of 20. Though the accuracy drops to 93%, the model still correctly classifies healthy and defective

responses according to its true ground class as is visible from the confusion chart.

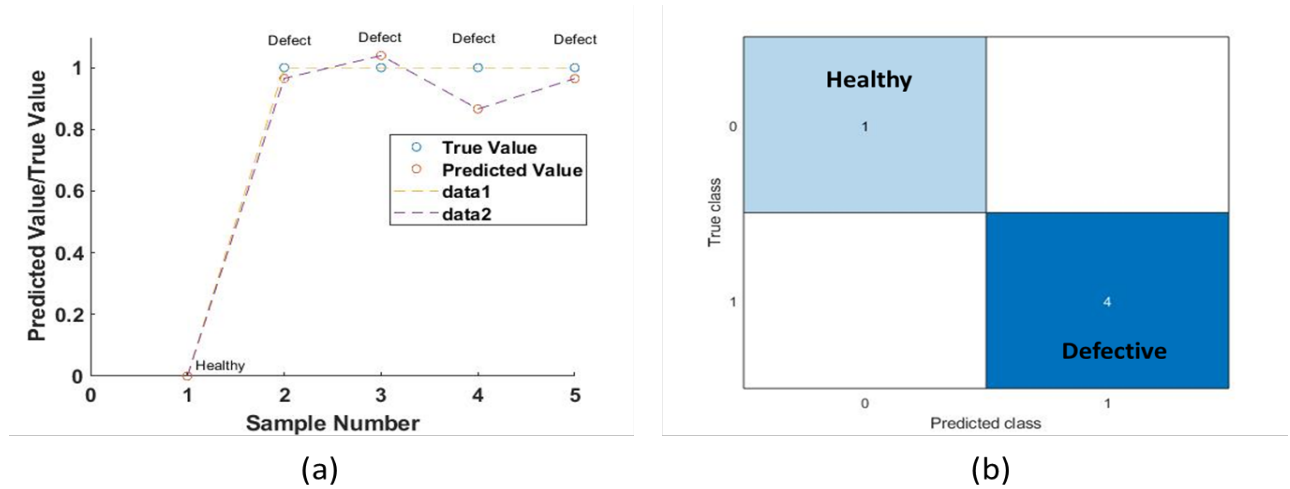


Figure 56. (a) Prediction results for the MLP network and (b) confusion plot for the predicted results

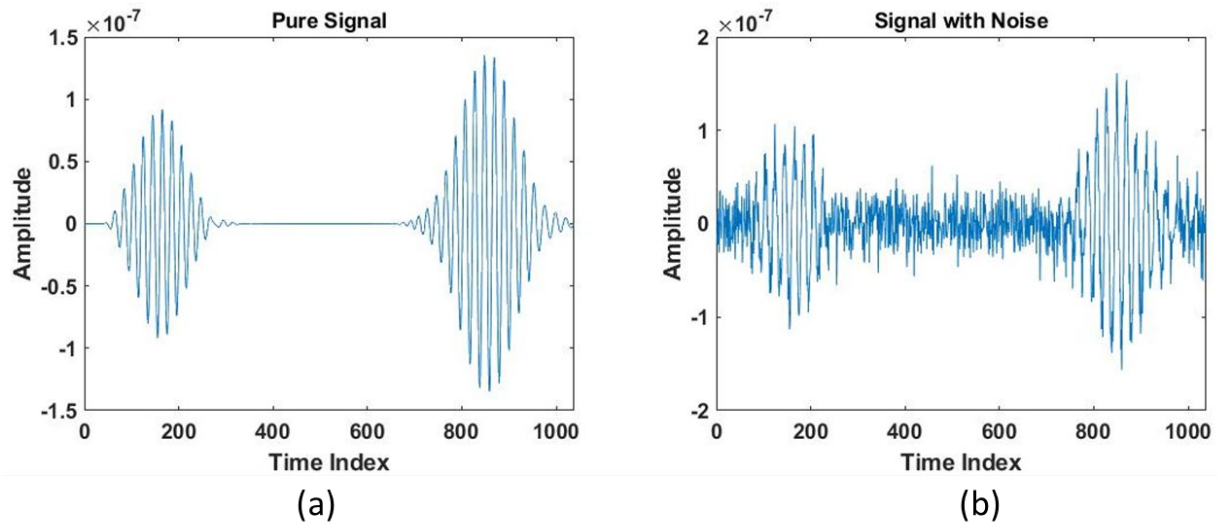


Figure 57. (a) Pure raw signal and (b) signal with additive white Gaussian noise (AWGN) added at SNR 5

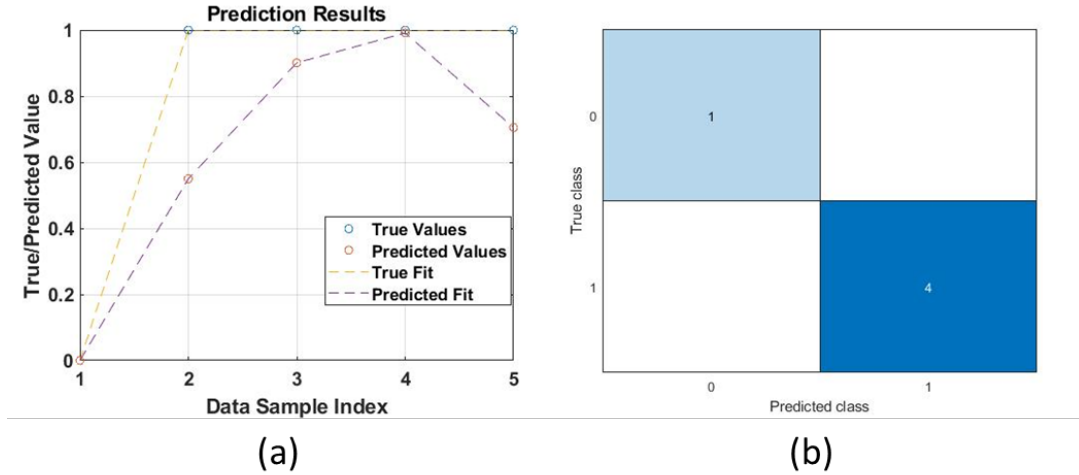


Figure 58. (a) Prediction results for the MLP network trained with a dataset with SNR 20 (b) confusion plot for the predicted results

Similarly, when the noise level is increased by maintaining a SNR level of 10, the performance further drops to 91% accuracy. Though it classifies the responses accurately, the absolute values the MLP is predicting for a defect are further away from 1, as is visible Figure 59. For a SNR level of 5, the accuracy further drops to 89%, and the network now falsely classifies a healthy response as a defective response as seen in Figure 60. But despite this error, the MLP network correctly classifies all defective responses correctly. Though this is not the best scenario, it is acceptable as when it comes to nondestructive evaluation, it is sometime acceptable to falsely classify healthy signals as long as defective signals are classified correctly. Thus, we are able to attain very good performance at SNR levels of 10 and 20, and acceptable performance at a SNR level of 5.

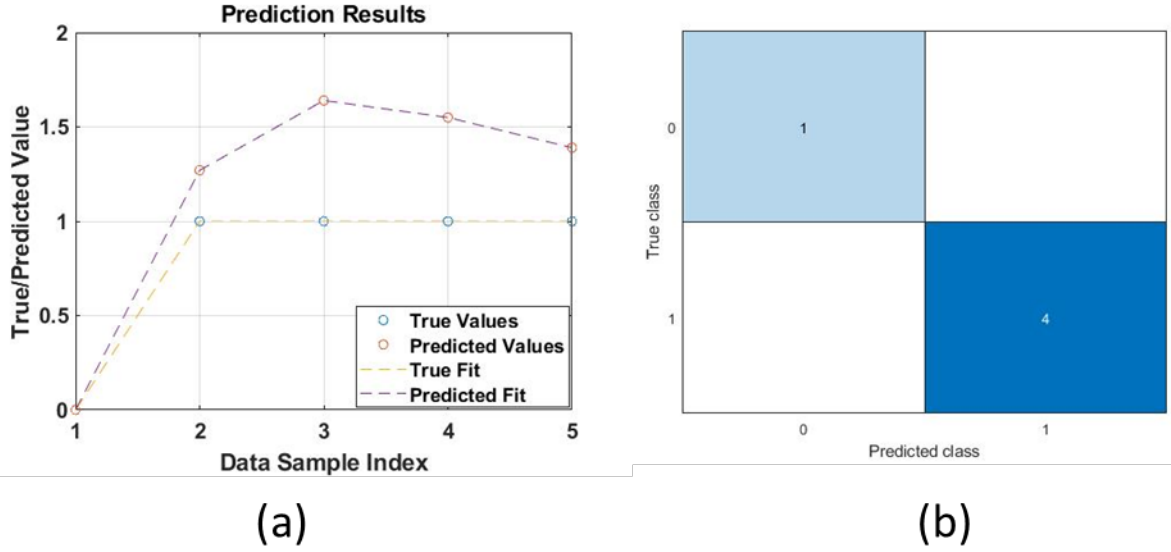


Figure 59. (a) Prediction results for the MLP network trained with a dataset with SNR 10 (b) confusion plot for the predicted results

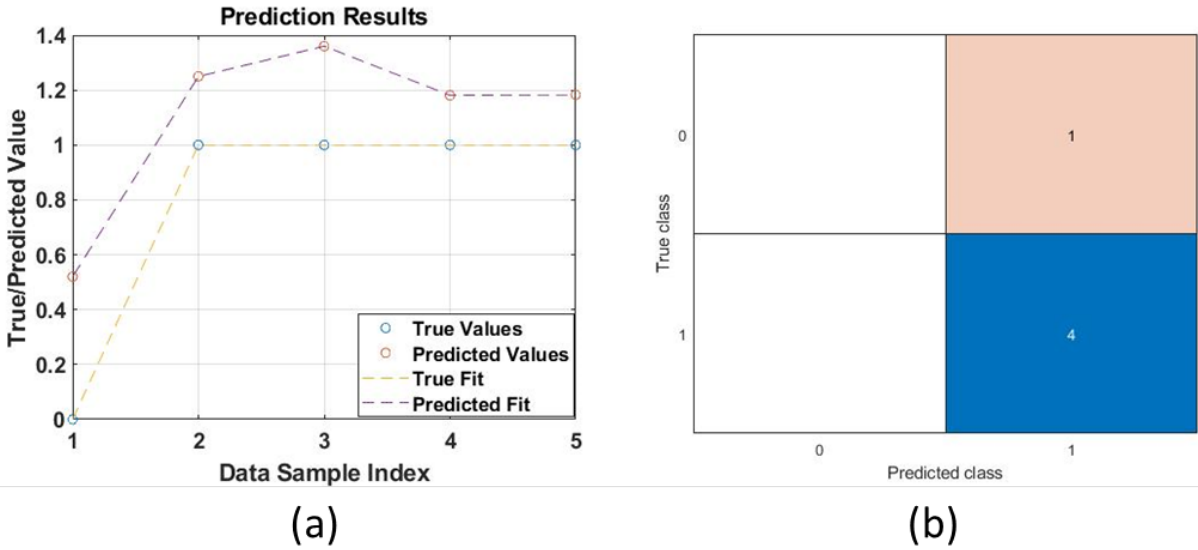


Figure 60. (a) Prediction results for the MLP network trained with a dataset with SNR 5 (b) confusion plot for the predicted results

4.3.4 1D-Convolutional Neural Network

The classifier network has been optimized by training it with signals added with different noise levels, thereby making the network robust and reliable. The next step is to develop 1D Convolutional Neural Network (1D-CNN) to help characterize the corrosion defects. 1D-CNN is known to extract inherent features from long time series data. Extensive research has gone into

using CNNs for defect localization and characterization in composites in the last few years. In the 1D-CNN network proposed in this study, the input will be the velocity time histories obtained, and the output will be the defect parameters such as defect depth, defect length, and number of corrosion defects in the pipe. One of the reasons for adopting 1D-CNN is that it trains faster than recurrent neural networks.

Sparse connections and parameter sharing are two important ideas in CNN, whereas, in a fully connected network (FCN), every neuron interacts with every other neuron. CNN helps to reduce the number of learnable parameters, which eventually saves memory and decreases the training time. CNNs are also very robust to external influences, and generally have been shown to perform well even when there is low level noise in the data. 1D-CNN works similar to a traditional CNN/2D-CNN, the only difference is that the inputs, kernels and feature maps are all in one dimension. Figure 61 shows the framework of the proposed 1D-CNN model. Please note the CNNs are data hungry, and with a current dataset of size 150 samples, it is incredibly hard to attain acceptable performance. Therefore, the current model shown is being fine-tuned using the small dataset, while simultaneously more data is being created using data augmentation techniques along with FEM simulations to populate the dataset.

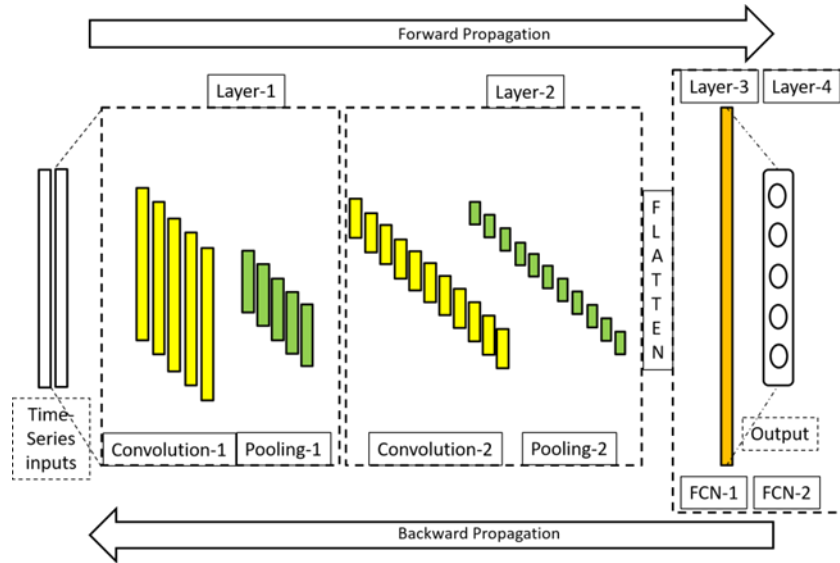


Figure 61. Architecture of the proposed 1D-CNN

4.3.5 Conclusions

In this Section, different defect features were established to be used in an overall learning paradigm

that leverages the best of Machine Learning based feature engineering algorithms, and Deep Learning based Convolutional Neural Networks (CNNs). A Multi Cross Entropy Analysis (MCEA) was conducted, and a Damage Index (DI) based on the similarity number was established to quantify the defect depths, and the number of corrosion defects. Also, Hilbert Transform (HT) analysis was conducted to establish another feature of interest. The multi-layer perception (MLP) classifier was further improved by training based on the data with different noise levels. It was seen that even at very high noise levels (such as SNR of 5), the performance of the networks was acceptable. Furthermore, a 1D-CNN network was proposed for characterization of corrosion defects, which currently is being fine-tuned.

4.4 Ultrasonic Imaging Methods

For the same corrosion samples for ECA imaging and data analysis, MSU team also acquired the immersive UT imaging data, which are shown in the figures below, compared with the IFM data obtained by Akron team.

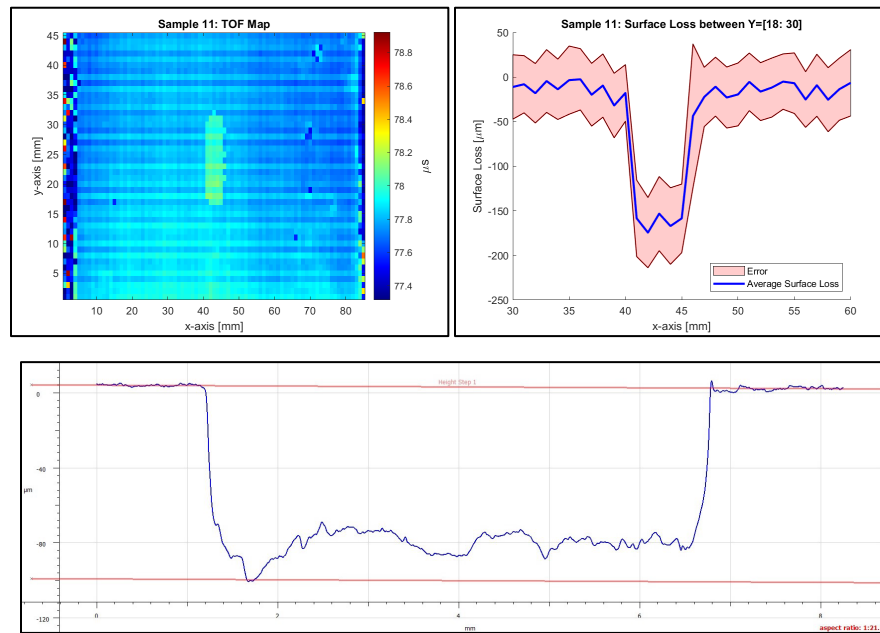


Figure 62. Sample 11 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

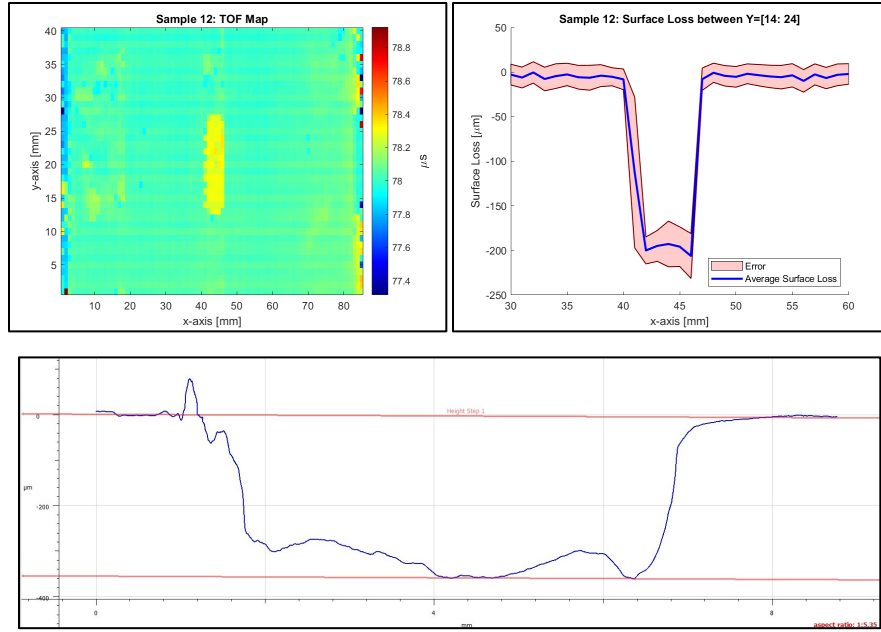


Figure 63. Sample 12 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

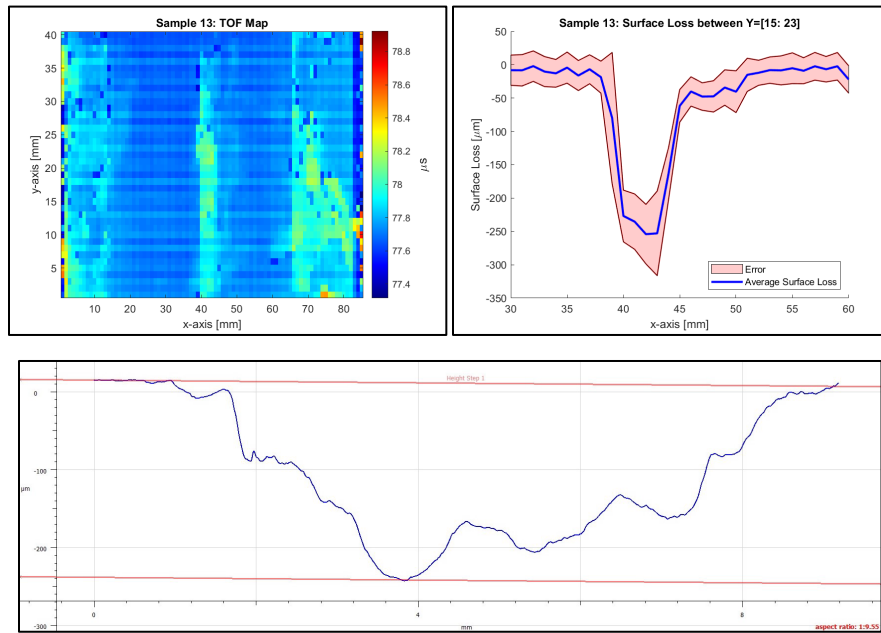


Figure 64. Sample 13 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

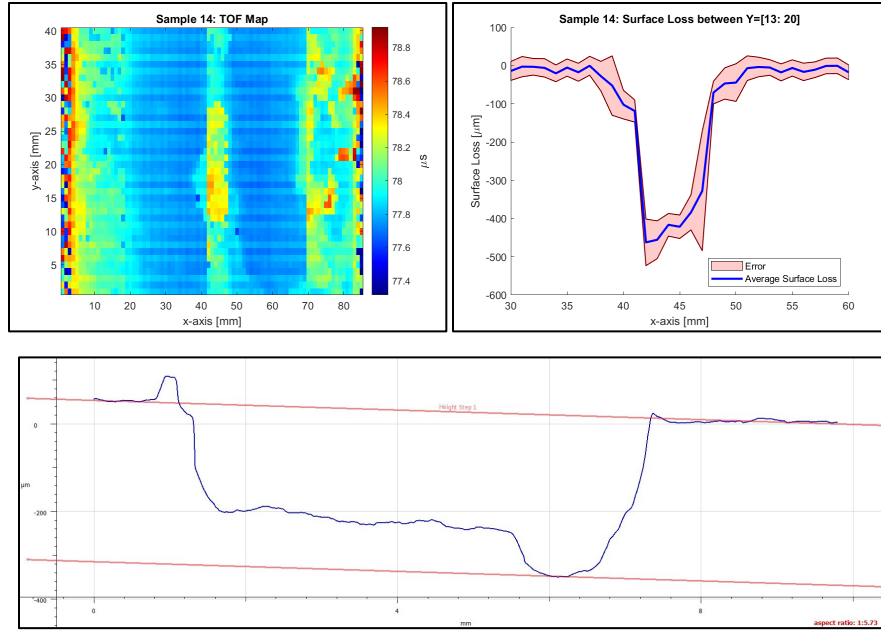


Figure 65. Sample 14 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

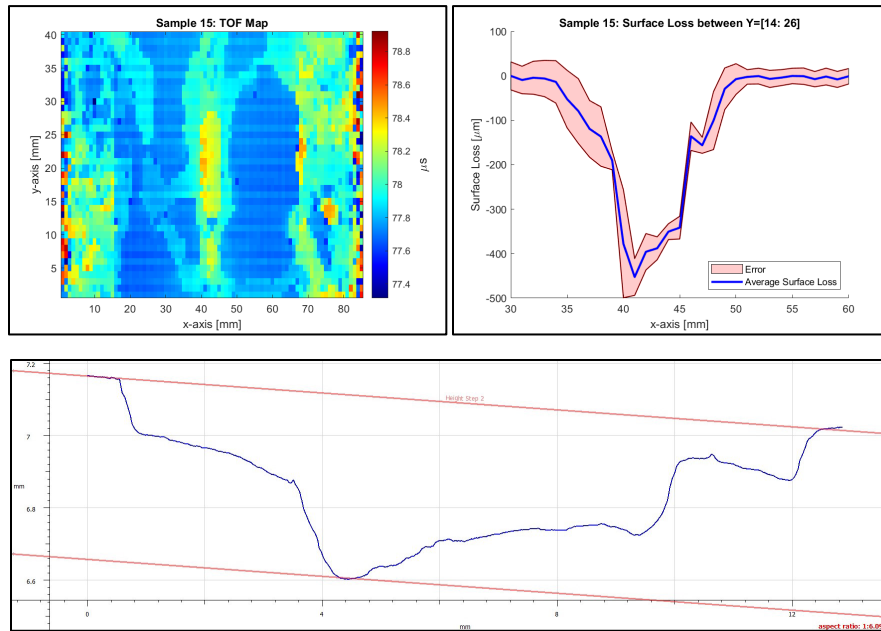


Figure 66. Sample 15 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

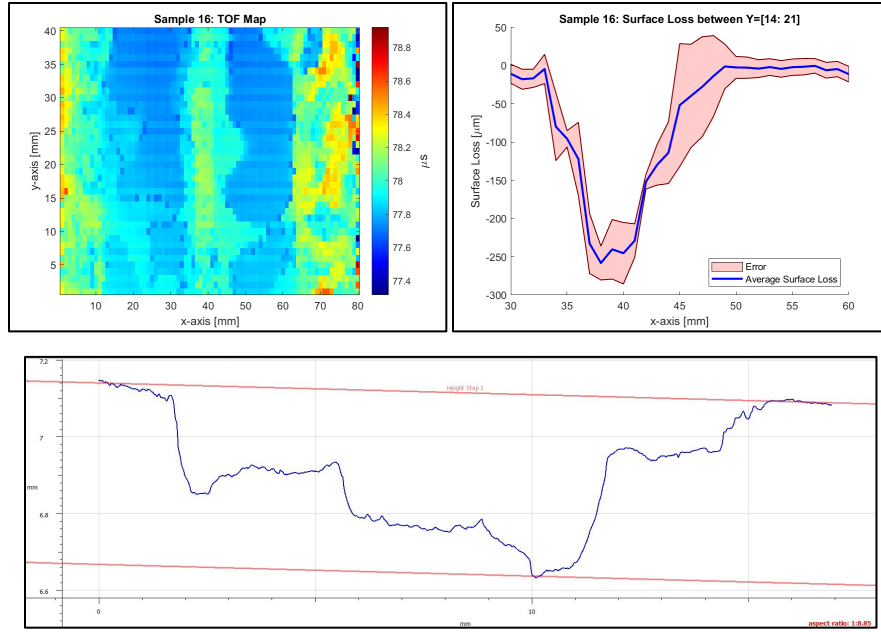


Figure 67. Sample 16 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

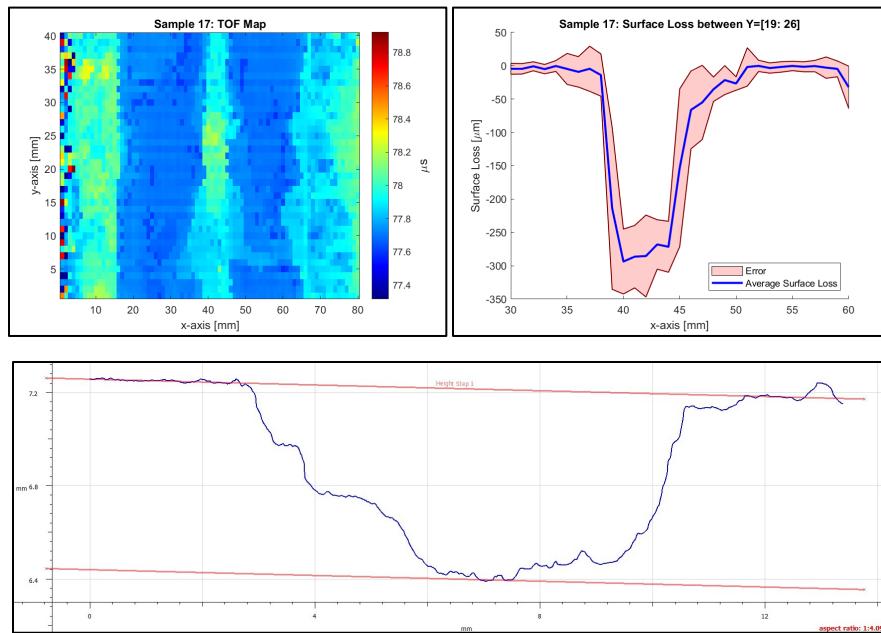


Figure 68. Sample 17 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

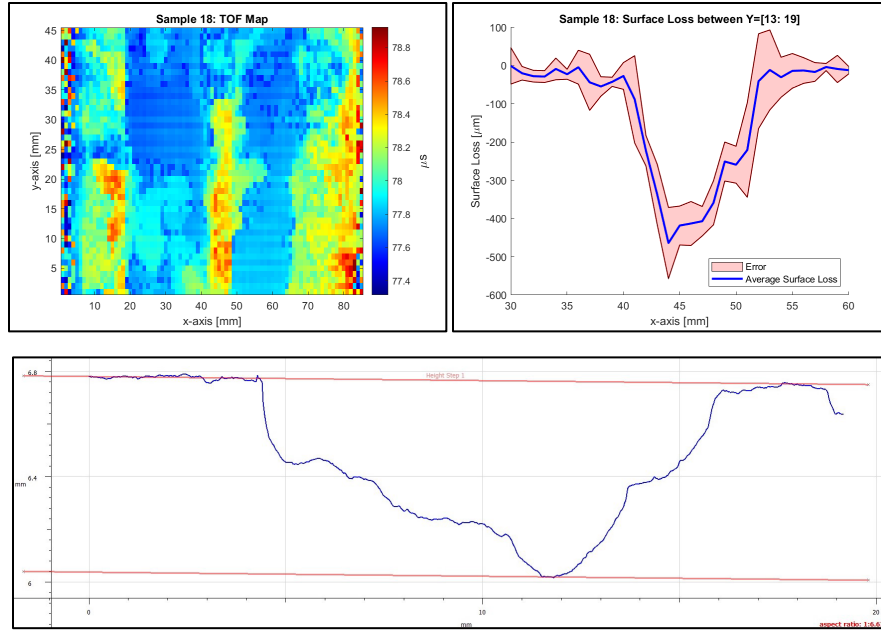


Figure 69. Sample 18 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

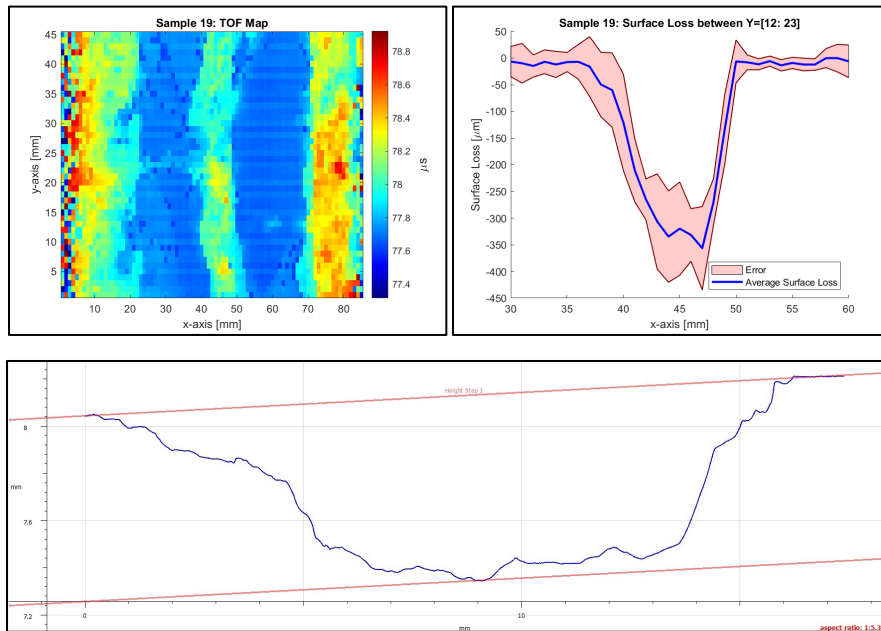


Figure 70. Sample 19 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

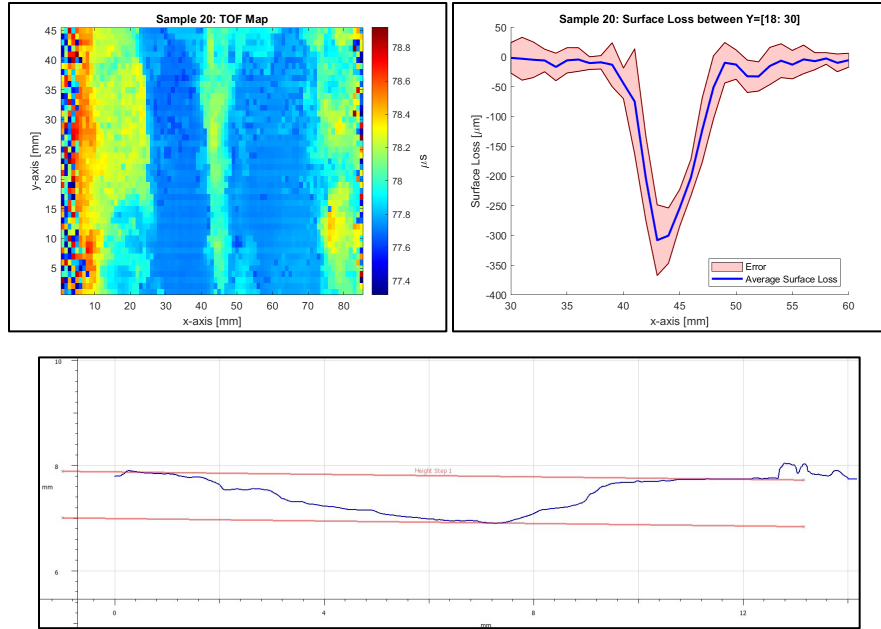


Figure 71. Sample 20 UT results: (a) 2D imaging/TOF map, (b) Estimated corrosion profile with uncertainties using UT, and (c) IFM estimated corrosion profile as ground truth.

It can be seen from both ECA and UT results that NDE results correlate with the IFM line data (corrosion “ground truth” profile) well. Further research will be conducted in multi-modal data fusion that combines ECA, UT and IFM data at both measured data-level and feature-level. The time-dependent corrosion process and damage characterization will be performed with uncertainties quantification, and passed to the probabilistic modeling for failure pressure prediction.

5 Objective 3: Probabilistic models of failure pressure prediction

5.1 Models for Pipeline with Single Corrosion Defect

5.1.1 Background

In practice, the remaining strength of pipelines with single corrosion defects has been popularly assessed using the ASME B31G [5], Modified ASME B31G [5], or RSTRENG [5] methods. These methods were developed based on the NG-18 equation proposed by the Battle Memorial Institute. More models based on the NG 18 equations were also developed later such as SHELL92 [6], RPA[7], and DNV RP-F101 [8]. On the other hand, other researchers (such as Netto et al [9]; Mustaffa and Van Gelder [10]; Wang and Zarghamee [11]) proposed failure prediction models based on Buckingham's π theorem, which is a mathematical approach that allows the formation of dimensionless parameters consisting of various possible influencing parameters and then uses the dimensionless parameters to predict a quantity of interest [10]. Zhu and Leis [12] proposed a prediction model using the strain-hardening behavior of pipe materials. A comprehensive review of the existing prediction models has been reviewed in [13]. Note that all existing prediction models mentioned here were developed for pipelines under normal range temperatures. Separate prediction models need to be developed for pipelines under extremely low temperatures, which significantly change material properties (such as toughness) and thus pipeline capacity. For instance, Chen et al. [14] have developed a burst pressure model for corroded hydrogen storage pipeline at extremely low temperature.

To appropriately incorporate the underlying uncertainties in the pipeline risk management, the model error of the remaining strength prediction model needs to be assessed. The model error reflects the bias and extent of variability in the prediction model; thus, it can be used to measure the performance of a prediction model. Several studies were conducted to compare the performance of some existing prediction models based on model error. For example, Zhou and Huang [15] performed model error comparison for 8 models (i.e., ASME B31G [5], Modified B31G [5], CPS [14], CSA [15], DNV RP-F101 for single defect (Part B) [8], PCORRC [16], RSTRENG [5] and SHELL92 [6]) based on 150 full-scale burst test results for pipelines with isolated real corrosion defects. They found that RSTRENG had the best performance overall, and all the models became more conservative for long corrosion defects. A similar model performance comparison was also done by Amaya-Gómez et al [13] but based on three categories of material toughness. The results showed that the performance of the models varied for different levels of

material toughness, implying that some models have applicability limits for specific material toughness. This is not surprising, as many of the prediction models were developed specifically for certain types of grade. Other studies compared prediction models based on the probability of failure but without incorporating model errors in the analysis. For instance, Hasan et al. [17] and Amaya-Gómez et al. [13] compared the failure probability results from different prediction models by only considering uncertainties in the operating conditions and the pipeline material and geometry properties. Based on the comparison study by Amaya-Gómez et al. [13], most of the existing prediction models are shown to be conservative with different levels of conservatism. Conservative methods may be good for design in terms of safety; but when they are applied in the risk management, it leads to unnecessary costs associated with inspection, repairs, and maintenance. Therefore, an accurate prediction model with less bias and variability is needed for decision-makings in a cost-effective risk management.

5.1.2 Existing prediction models

The assessment methods of the remaining strength of corroded pipelines can be classified into three levels based on the available level of information and the degree of precision required [17]. A Level-1 assessment method uses the maximum depth and projected axial length of a defect to evaluate the remaining strength. It does not consider the shape of the defect and defect interactions due to a cluster of corrosion defects; and is generally conservative. A Level-2 assessment method evaluates the remaining strength considering the possibility of interaction between defects or the impact of the defect shape. A Level 3 assessment refers to nonlinear finite element (FE) analysis method, which requires the maximum information of material properties and defect configuration; and this level assessment usually can provide accurate failure predictions with an error of around 5%.

In the literature, most of the prediction models for the remaining strength of a corroded pipeline are Level-1 assessment, which only considers maximum length and depth of the corrosion defect. Those models are applicable for isolated defects when defect interactions can be ignored, which is the focus of this study. In this section, the existing prediction models are reviewed first. The formulations of 24 existing prediction models are summarized in Appendix B. Table 3 is modified based on Amaya-Gómez et al. [11] and provides a comparison of these existing models in terms of application restrictions (i.e., yield strength level, pipeline grade, and defect geometry) and prediction performance based on the analysis conducted in this study. Note that the application

restrictions of most models are suggested during the model development. In this study, the 24 existing models are grouped into six groups based on how the model is developed, and each model is given a short name based on its group number (as shown in Table 3). The first group of models (G1) is developed based on the NG-18 equation [27]; the second group (G2) is based on Buckingham's π theorem; the third group (G3) is constituted of models based on FE models with a plastic collapse failure criterion called PCCOR; the fourth group (G4) has the models that use a stress concentration factor in the model formulation and neglect the defect length; the fifth group (G5) is formed by models that use the strain-hardening behavior of pipelines; and the last group (G6) consists of models that are developed using other approaches. Also, Table 3 gives the defect length limit beyond which a defect is classified as a long defect in some models.

Table 3. Comparison of 24 existing pressure failure prediction models

Group	Model	Grade	Other restrictions	Long defect limit	Performance comparison within the group ^a
G1: Models based on NG-18	G1-1: ASME B31G Original [3]	Below X56	$0.1 \leq d/t \leq 0.8$	$l^2/Dt > 20$	All the models have similar performance and G1-7 has the best performance overall
	G1-2: Modified B31G [3]	Below X65	$0.1 \leq d/t \leq 0.8$	$l^2/Dt > 50$	
	G1-3: SHELL92 [4]	-	$d/t \leq 0.85$	-	
	G1-4: RPA [5]	Below X65	$0.1 \leq d/t \leq 0.8$	$l^2/Dt > 20$	
	G1-5: RSTRENG Effective Area [3]	-	$0.1 \leq d/t \leq 0.8$	$l^2/Dt > 50$	
	G1-6: CSA Z662 [15]	Below X65	-	$l^2/Dt > 50$	
	G1-7: DNV RP-F101 [6]	Except X80	-	-	
	G1-8: Fitnet FFS [19]	-	-	$l^2/Dt > 20$	
	G1-9: Phan et al Modified NG-18 [20]	-	-	-	
G2: Models based on Buckingham's π theorem	G2-10: Netto et al [7]	X52 – X77	$0.1 \leq d/t \leq 0.8, l/D \leq 1.5, w/D \geq 0.0785$	-	G2-11 overestimates the burst pressure and has the worst performance overall but can be suitable for shallow defects.
	G2-11: Mustaffa & van Gelder [8]	-	$d/t \leq 0.3, l/D \leq 0.2, w/t > 0.5$	-	
	G2-12: Netto et al [21]	-	Shallow, moderately deep, and deep-narrow defects	-	
	G2-13: Wang & Zarghamee [9]	-	-	-	
	G2-14: Phan et al. Modified - Netto et al. [20]	-	-	-	
G3: PCORRC models	G3-15: PCORRC [16]	-	-	-	G3-15 has the best performance overall
	G3-16: Modified PCORRC [22]	X65 – X70	-	-	
G4: RAM PIPE Requal models	G4-17: Original Ram Pipe Requal [23]	-	-	-	G4 models do not perform well in general and have large variability, but G4-17 can be suitable for shallow defects.
	G4-18: Modified Ram Pipe Requal [23]	-	-	-	
G5: Models using strain-hardening	G5-19: Zhu & Leis [10]	Grade B and X80 Only	-	-	Both models have good performance, particularly G5-19
	G5-20: Zhu - X65 [24]	X65 – X80	-	-	
G6: Other approaches	G6-21: Choi et al. [25]	X65	-	$l^2/Dt > 18$	G6-22 and G6-23 perform the best overall; particularly G6-23 is good for thin pipe thickness and medium and long defects
	G6-22: Chen et al. [26]	X80 – X90	-	$l^2/Dt > 25$	
	G6-23: CUP [27]	X46 – X60	-	-	
	G6-24: Phan et al. - Modified Gajdoš et al. [20]	-	-	-	

^a The performance is based on three levels of σ_y , D/t , d/t , and l^2/Dt

5.1.3 Data Collection

To evaluate the performance of the existing prediction models and to develop a robust and accurate model later, comprehensive failure pressure data are needed. The database established in this study consists of the data directly collected from literature and additional numerical data obtained from FE analysis conducted in this study.

Data collected from literature

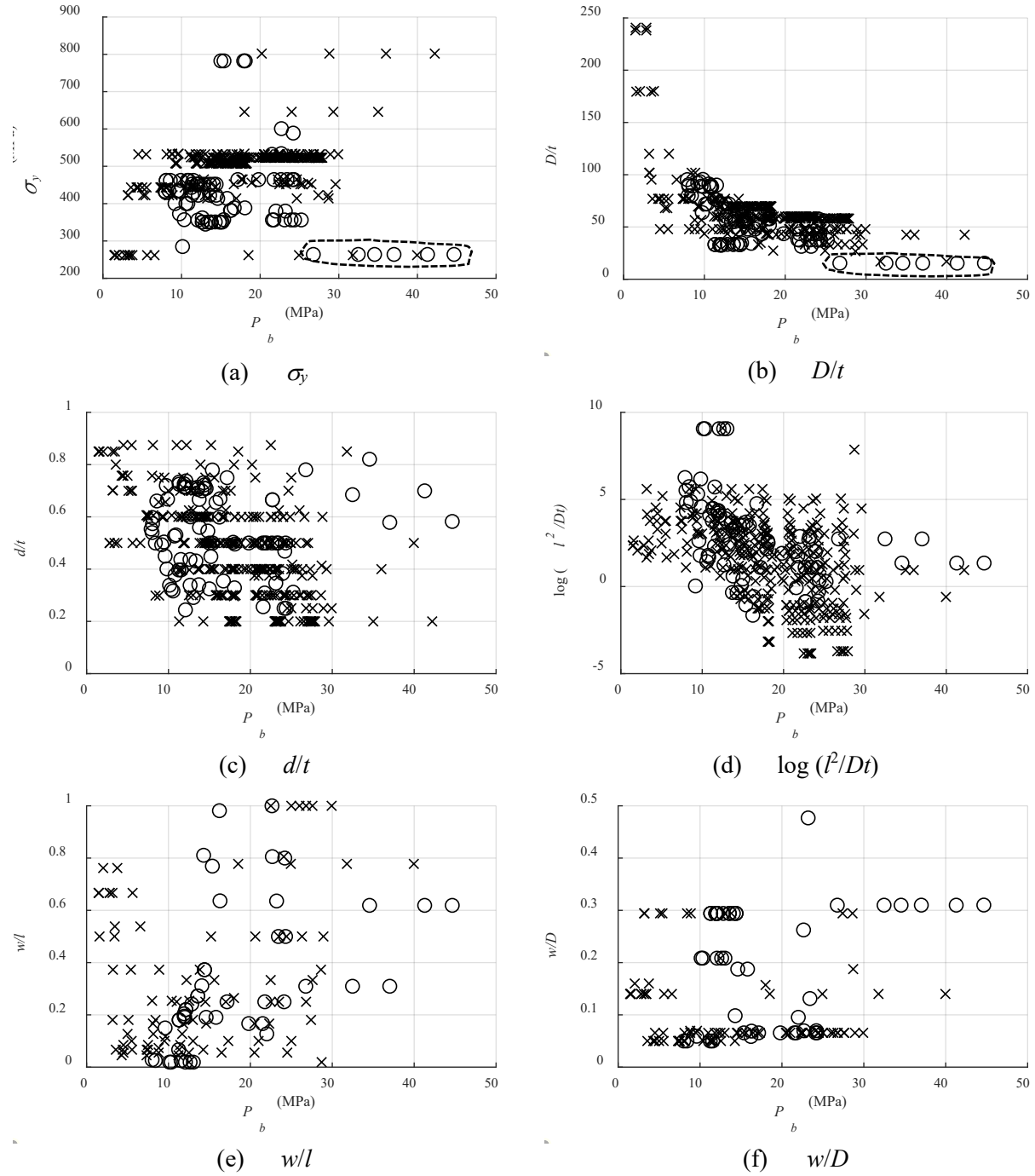
A total of 401 different burst test results are collected from literature, out of which 83 are laboratory experimental burst tests and 318 are FE simulations. In all these experimental tests or simulations, single defects are introduced to the external surface of the pipeline. Table 4 summarizes the overall ranges of six important quantities in the whole dataset (i.e., collected data and numerical data generated in this study): yield strength (σ_y), ratio of pipe diameter to pipe thickness (D/t), ratio of defect depth to pipe thickness (d/t), ratio of defect length squared to the multiplication of pipe diameter and thickness (l^2/Dt), ratio of defect width to defect length (w/l), and ratio of defect width to pipe diameter (w/D). These six quantities are listed out here because of their potential impact on failure pressure prediction, and they have been used in the failure pressure prediction as shown in Appendix B. Except σ_y and D/t , the other four quantities are related to defect geometry. Note that typically the information of w is not recorded. In the data collected from literature, 264 cases (more than a half of the total cases) do not contain w values. The overall ranges of w/l and w/D in Table 4 are based on the cases that have w values.

Table 4. Data range of six important quantities

Quantity	Overall Range	Level 1 Range (number of data)	Level 2 Range (number of data)	Level 3 Range (number of data)
σ_y (MPa)	[262 802]	[262 433] (88)	(433 508] (139)	(508 802] (206)
D/t	[15.34 240.63]	[10 40] (59)	(40 60] (235)	(60 250] (139)
d/t	[0.10 0.87]	[0 0.33] (131)	(0.33 0.67] (232)	(0.67 1] (70)
l^2/Dt	[0.018 8967.97]	[0.018 18]	(18 50]	(50 8967.97]
$\log(l^2/Dt)$	[-4.00 9.10]	[-4.00 2.890] (287)	(2.89 3.912] (71)	(3.91 9.10] (75)
w/l	[0.02 10.92]	-	-	-
w/D	[0.05 0.48]	-	-	-

To provide a better idea of how these six important quantities scatter over the database

collected from literature, Figure 72 shows the scatter plots of measured failure pressure (P_b) vs. the six quantities. In these plots, the circle and cross markers refer to the data obtained from experimental and numerical burst tests, respectively. Figure 72(a) displays the scatter plot of σ_y over P_b . As expected, the result indicates that in general higher yield strength leads to higher burst pressure, except a few cases circled by the dotted lines. It turns out these outline cases are the ones with low D/t values (referring to very thick-wall pipes), corresponding to the ones circled by the dotted lines in the scatter plot of D/t over P_b in Figure 72(b). Figure 72(b), (c) and (d) show the scatter plots of D/t , d/t , $\log(l^2/Dt)$ over P_b , respectively. A negative correlation is all observed for each of these three plots, indicating that the pipe diameter to pipe thickness ratio, the defect depth, and the defect length have negative effect on the failure pressure, as expected. However, no clear trend is observed in both Figure 72 (e) and (f), indicating that the impact of defect width w may be insignificant on the burst pressure. Note that Figure 72(e) and (f) are plotted using only the data that have w information.



○ Experimental × Numerical

Figure 72. Scatter plots of burst pressure (P_b) vs. selected quantities

Additional numerical data

To complement the data collected from literature, FE analysis is conducted to generate additional cases. Software ANSYS or ABAQUS has been widely used in research to obtain failure pressure of pipelines with defects. In this study, FE models are developed in ABAQUS. For computational efficiency, the corrosion defect is modeled as a rectangular shape. In addition, thanks to symmetry, only a quarter of the pipe with appropriate boundary conditions is modeled, as shown in Figure 73 to further reduce the computational cost. ABAQUS Statics-General procedure is used for the analysis, and the burst pressure is determined when the von Mises stress at any point of the defect area reaches the ultimate tensile strength of pipe steel [25].

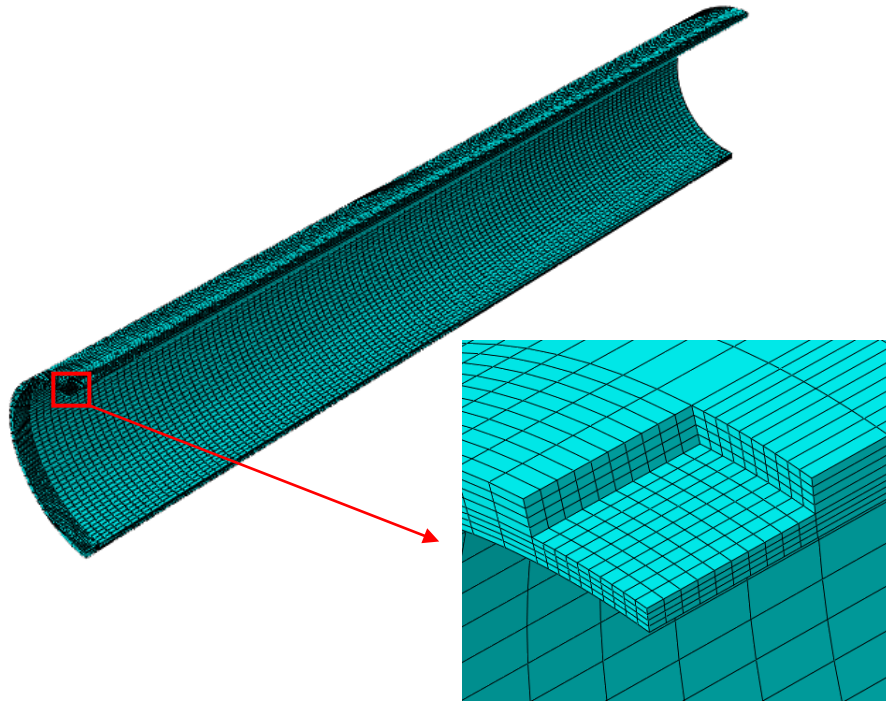


Figure 73. A quarter of a corroded pipe modeled in ABAQUS

Before the FE analysis is used to generate new cases, a few experimental burst test results from the literature are selected for the model validation. Table 5 shows the pipe material and geometry properties and defect geometry of the selected cases, and it also compares the failure pressure data reported in the literature ($P_{b,test}$) and the failure pressure obtained from the FE analysis ($P_{b,FE}$) conducted in this study. The cases are selected with the intention to cover a wide range of pipe grade: from AISI1020 Mild (low grade) to X100 (high grade). Please note that for

the selected cases where the information of w is missing, w is assumed, since the effect of w on the burst pressure is found to be insignificant. In this study, w is assumed to be $0.05\pi D$ based on an assumption used in Choi et al. [24]. As shown in Table 5, the error percentages, $(P_{b,\text{test}} - P_{b,\text{FE}})/P_b$, are all within 10% except only one case (whose error is about 16%). This result validates the FE analysis.

With the validated FE models, a total of 32 new additional cases are generated for the burst numerical analysis and the results are added to the database. The new cases are designed to cover the regions where the data collected from literature are scarce.

Figure 74 shows the new FEM cases (marked as stars) and the cases collected from literature (marked as circles for experimental cases and crosses for FE cases) in terms of the four important quantities (σ_y , D/t , d/t , and $\log(l^2/Dt)$). The pipeline properties and defect geometries and FE results ($P_{b,\text{FE}}$) of these 32 new cases are provided in Table 6.

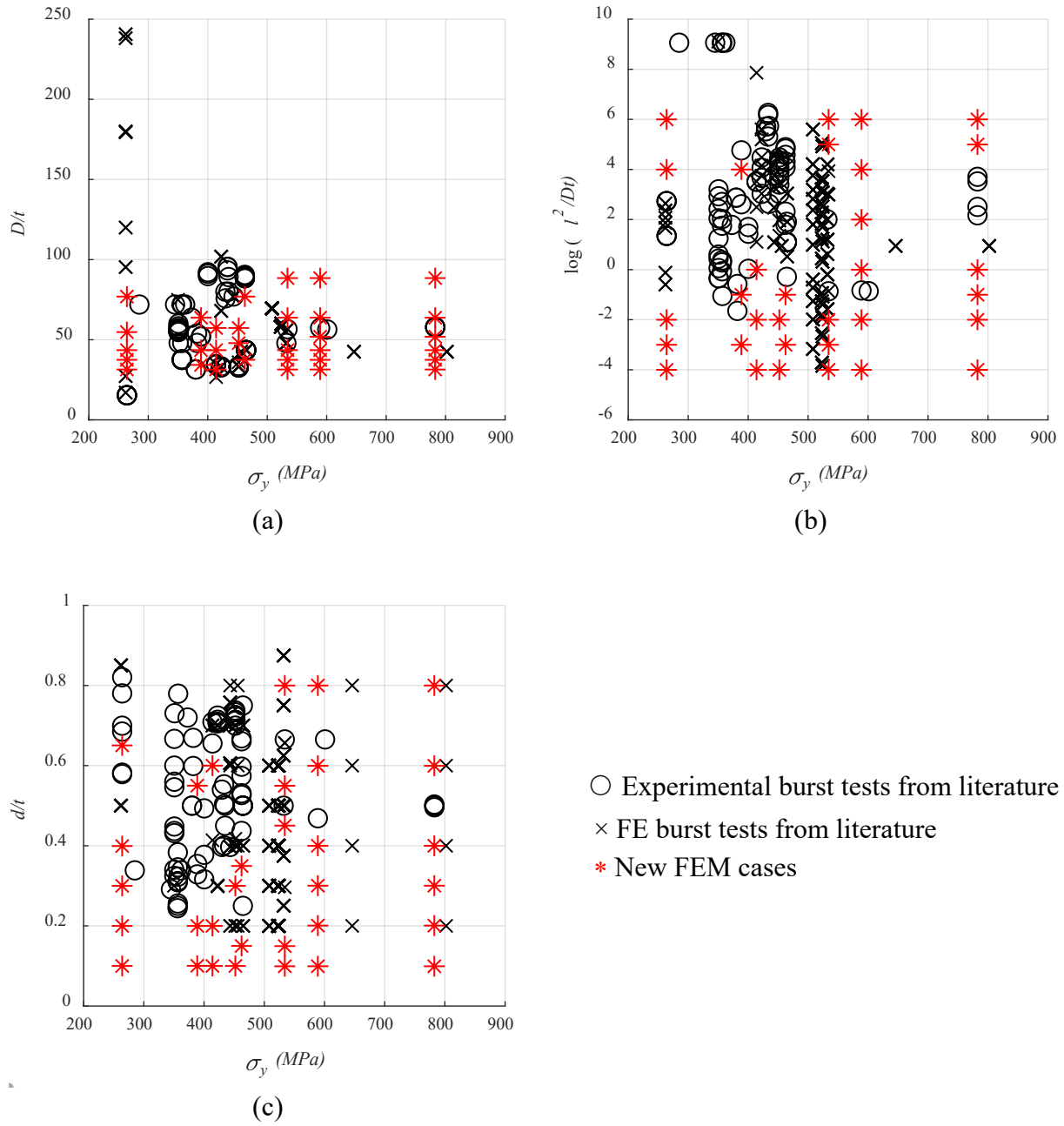


Figure 74. Scatter plots of selected quantities vs. yield stress (σ_y)

Table 5. Experimental testing cases selected from literature for validating the FE models

Reference	Grade	Diameter D (mm)	Thickness t (mm)	σ_y (MPa)	σ_u (MPa)	d (mm)	l (mm)	w (mm)	$P_{b,test}$ (MPa)	$P_{b,FE}$ (MPa)	Error (%)
[7]	AISI1020 Mild	42	2.73	264	392	1.58	42.00	13.00	37.02	36.66	0.97
[7]	AISI1020 Mild	42	2.73	264	392	2.24	21.00	13.00	34.55	33.42	3.28
[29]	X42	274	4.93	351	454	1.60	45.72	43.02a	14.99	16.24	8.33
[29]	X42	274	4.57	351	454	2.74	66.04	43.11a	12.67	13.03	-2.86
[29]	X46	323	8.64	356	469	2.16	63.50	50.79a	24.37	26.71	-9.59
[29]	X46	864	9.47	400	508	3.00	185.42	135.65a	10.56	11.18	-5.89
[29]	X52	273	5.26	389	502	1.73	139.70	42.89a	18.06	17.76	1.67
[29]	X52	612	6.40	433	535	2.57	1371.60	96.05a	9.81	8.20	16.44
[29]	X55	508	5.64	462	587	2.46	170.18	79.8a	11.51	11.59	-0.75
[29]	X55	507	5.74	462	587	3.02	132.08	79.60a	10.73	11.63	-8.44
[30]	X60	324	9.74	452	542	7.14	528	95.3	11.3	10.71	5.20
[30]	X60	508	14.8	414	600	9.7	500	95.3	15.8	16.24	-2.77
[31]	X65	762	17.5	465	564	8.75	300	50	19.8	20.08	-1.43
[31]	X65	762	17.5	465	564	8.75	100	50	24.3	25.85	-6.38
[22]	X70	762	15.9	532	627	7.95	300	50	21.5	20.62	4.08
[32]	X80	459	8.1	534	661	5.39	39.6	31.9	22.68	22.25	1.92
[33]	X80	459	8.00	589	731	3.75	40.00	32	24.20	25.85	-6.82
[34]	X100	1321	22.81	782	803	11.31	608.05	207.47a	18.10	18.64	-2.99
[34]	X100	1321	22.81	782	803	11.41	1108.13	207.47a	15.40	16.92	-9.93

^a w is assumed to be $0.05\pi D$ based on an assumption from [25]

Table 6. New FE models cases

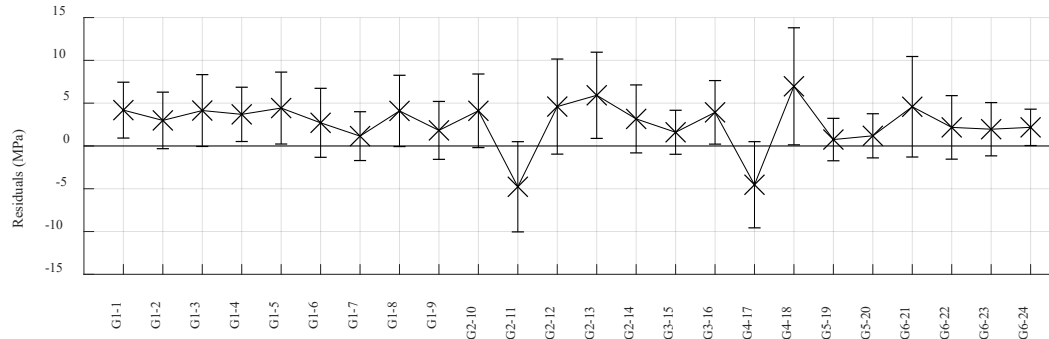
Grade	Diameter D (mm)	Thickness t (mm)	σ_y (MPa)	σ_u (MPa)	d (mm)	l (mm)	w (mm)	$P_{b,FE}$ (MPa)
AISI 1020 Mild	508	6.6	264	392	0.66	7.84	79.80	10.60
AISI 1020 Mild	274	5	264	392	1.00	8.26	43.04	14.94
AISI 1020 Mild	762	17.5	264	392	5.25	42.48	119.69	18.64
AISI 1020 Mild	324	8.64	264	392	3.45	390.86	50.89	15.84
AISI 1020 Mild	324	10.3	264	392	6.70	1160.31	50.89	10.67
X52	324	5.08	389	502	0.51	9.05	50.89	17.13
X52	762	17.5	389	502	3.50	70.04	119.69	24.97
X52	508	14.8	389	502	8.14	640.70	79.80	17.42
X60	459	8	414	600	0.80	8.20	72.10	22.88
X60	762	17.5	414	600	3.50	42.48	119.69	29.36
X60	324	10.3	414	600	6.18	57.77	50.89	34.50
X60	459	8	452	542	0.80	8.20	72.10	21.25
X60	762	15.9	452	542	4.77	40.49	119.69	24.23
X55	508	6.6	462	587	0.99	12.92	79.80	16.74
X55	324	8.64	462	587	3.02	32.08	50.89	32.34
X80	508	5.74	534	661	0.57	7.31	79.80	16.43
X80	324	5.08	534	661	0.76	9.05	50.89	22.81
X80	762	17.5	534	661	7.88	42.48	119.69	32.34
X80	324	8.64	534	661	4.75	644.01	50.89	19.19
X80	324	10.3	534	661	8.24	1160.31	50.89	10.19
X80	508	5.74	589	731	0.57	7.31	79.80	18.15
X80	324	5.08	589	731	1.02	14.92	50.89	24.96
X80	273	5.26	589	731	1.58	37.89	42.88	29.09
X80	762	17.5	589	731	7.00	313.90	119.69	28.01
X80	324	8.64	589	731	5.18	390.61	50.89	19.74
X80	324	10.3	589	731	8.24	1160.31	50.89	11.24
X100	508	5.74	782	803	0.57	7.31	79.80	20.51
X100	324	5.08	782	803	1.02	14.92	50.89	28.71
X100	273	5.26	782	803	1.58	22.98	42.88	34.57
X100	762	17.5	782	803	7.00	115.48	119.69	38.22
X100	324	8.64	782	803	5.18	644.01	50.89	20.61
X100	324	10.3	782	803	8.24	1160.31	50.89	12.34

Performance comparison of existing models

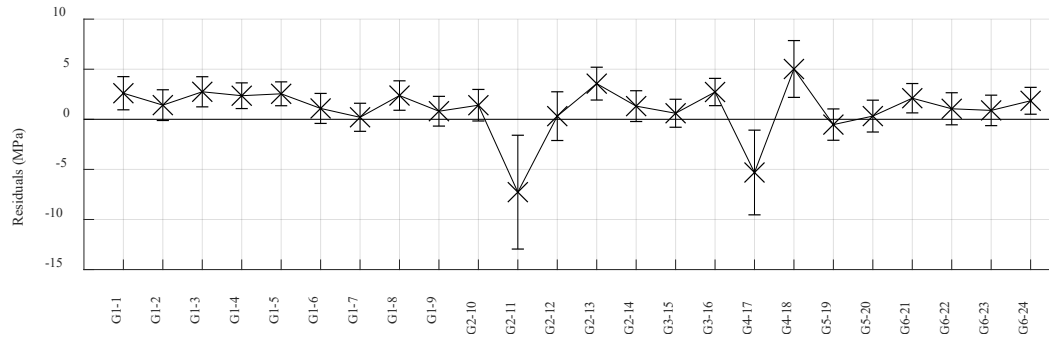
As shown in Table 3, some of the prediction models are developed for certain type of grade or defect geometry. Thus, it is more appropriate to compare the prediction models at different levels of grade and defect geometry. As indicated in Figure 72, the changes in four quantities (i.e., σ_y , D/t , d/t , and l^2/Dt) have shown tendency to change the failure pressure; thus, the performance of the existing models can be compared at different levels of these four quantities. In particular, σ_y reflects material strength, D/t describes the level of the relative pipe wall thickness, and d/t and l^2/Dt suggest the extent of corrosion. For simplicity, each quantity is classified into three predefined levels, as shown in Table 4. The three levels of σ_y correspond to low strength (level 1), moderate strength (level 2), and high strength (level 3). The three levels of l^2/Dt are based on the long defect limit suggested in the literature, where $l^2/Dt = 18$ suggested by Choi et al. [24] and $l^2/Dt = 50$ suggested in [3]. The three levels of d/t correspond to shallow depth, moderate depth, and deep depth, respectively. The three levels of D/t correspond to thick wall, moderate thick wall, and thin wall, respectively.

The performance of a prediction model can be quantified using mean (μ_{res}), standard deviation (σ_{res}), and mean squared error (MSE) of residuals (where residual refers to difference between the actual and the prediction values). In particular, MSE measures the combination of the prediction bias and variance. Figure 75 shows the performance comparison of the 24 models for three levels of σ_y , where the crosses refer to μ_{res} , the horizontal lines refer to $\mu_{res} \pm \sigma_{res}$, and solid dots are the MSE values. Note that the results for models G2-11, G2-12, G6-22, and G6-23 are calculated only using the cases that have the information of w , since these models require the w value.

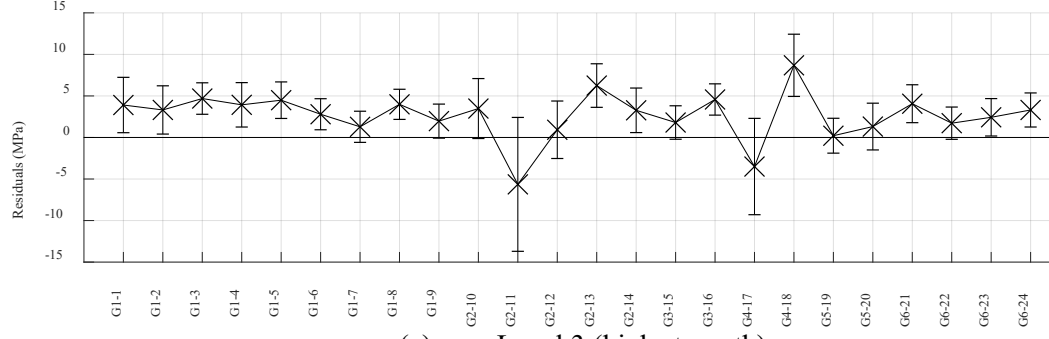
As shown in Figure 75, regardless the levels of σ_y , all prediction models (except G2-11 and G4-17) have positive μ_{res} (shown as cross markers above the horizontal line of zero residual, indicating they averagely underestimate the burst capacity. Figure 75 also indicates that most prediction models have smaller variability in the residuals for Level 2 (i.e., moderate strength) than for Levels 1 and 3 of σ_y , and the prediction variance is bigger for Level 1 of σ_y in general, shown in Figure 75(a). With MSE shown in Figure 75(d), one could also observe that most of models perform best in Level 2 and worst in Level 1. All these results indicate that the performance of each model changes from level to level.



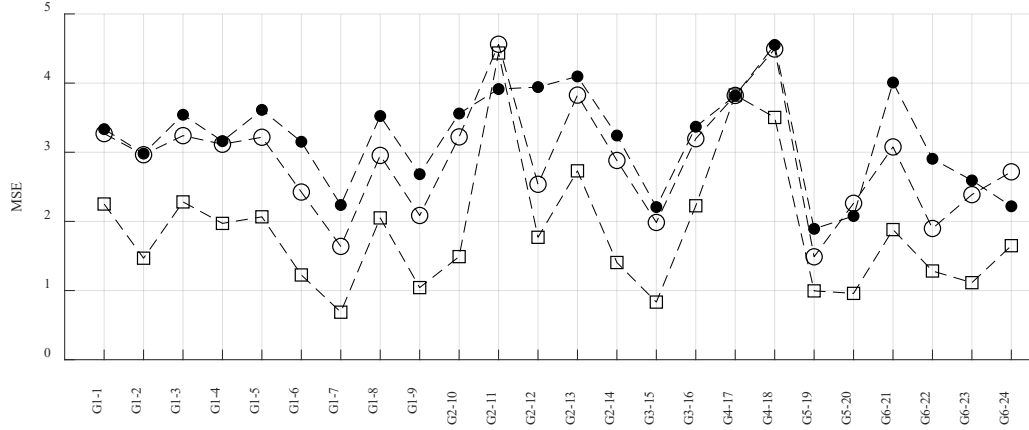
(a) σ_y – Level 1 (low strength)



(b) σ_y – Level 2 (moderate strength)



(c) σ_y – Level 3 (high strength)



(d) σ_y – MSE – All Levels

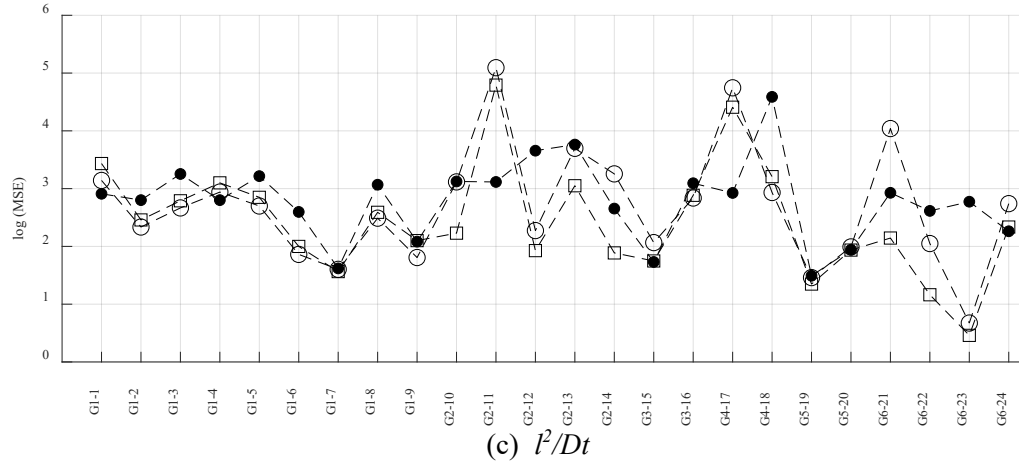
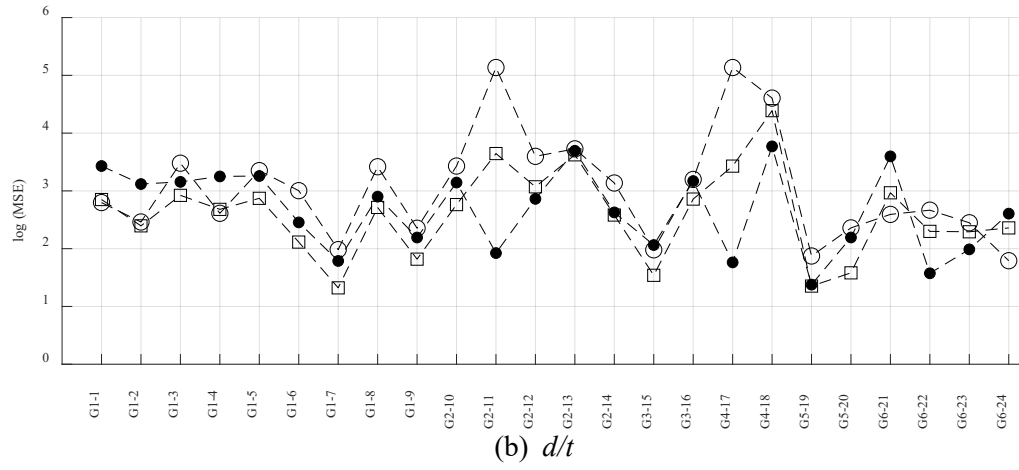
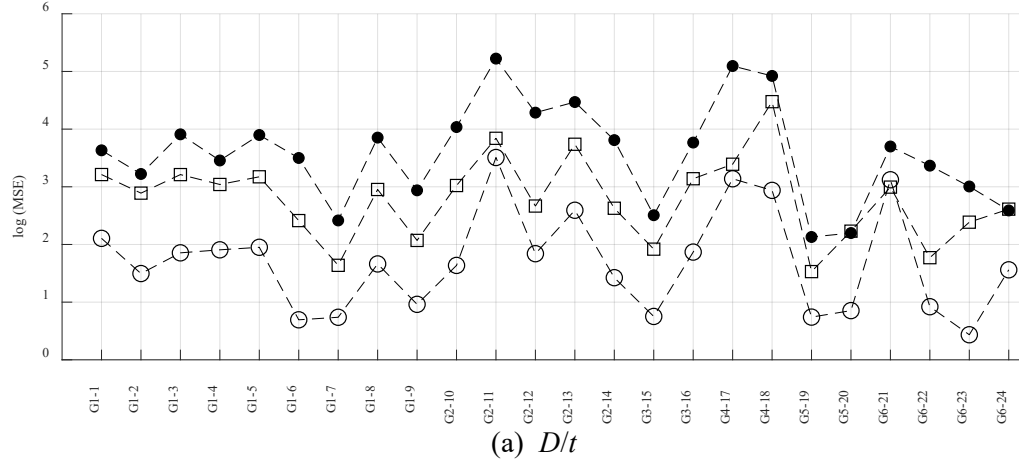
● Level 1 (low strength) □ Level 2 (moderate strength) ○ Level 3 (high strength)

Figure 75. Comparison of residual and MSE of each model for three levels of σ_y

Furthermore, different from all other models, G2-11 and G4-17 overestimate the burst capacity for all levels. Lastly, among G1 models, G1-7 has the best performance. Among G2 models, G2-11 model has the worst performance. Among the two G3 models, G3-15 is better and comparable with the best ones of all levels. While the two G4 models do not perform well at all, the two G5 models are the top models overall. Among the four G6 models, G6-22 and G6-23 perform best. The performance comparison of the existing models within their own groups is also summarized in Table 3.

Figure 76(a)-(c) show the comparisons of MSE of the 24 models for three levels of the other three quantities: D/t , d/t , and l^2/Dt , respectively. First, Figure 76(a) shows that most models have best performance for Level 3 of D/t (i.e., thin pipe thickness) and worst for Level 1 of D/t (i.e., thick pipe thickness). However, each model performs similarly among the three levels of d/t and l^2/Dt except a few cases, as shown in Figure 76(b) and Figure 76(c). For example, models G2-11 and G4-17 perform very differently for the three levels of d/t ; models G2-12, G4-18, and G6-23 perform much worse for Level 1 of l^2/Dt (i.e., short defects); model G6-23 has much better predictions for Levels 2 and 3 of l^2/Dt (i.e., medium and long defects). In addition, Zhou and Huang [12] found that the models considered in their study are more conservative for long defects, which is not observed in this study. The performance comparison of the existing models within their own groups based on levels of D/t , d/t , and l^2/Dt are also summarized in Table 3.

Furthermore, models are compared within their groups based on Figure 75 and Figure 76. The G1 models that are based on the NG-18 equations have similar prediction performance, and generally are pretty good models. Group G2 models based on Buckingham's π theorem have significantly different performance. The G4 models (that are developed by using the stress concentration factor in the model formulation and neglecting the defect length) have the worst performance overall, indicating that the impact of the defect length cannot be neglected. The G5 models based on the strain-hardening behavior of pipelines perform the best overall.



● Level 1 (low strength) □ Level 2 (moderate strength) ○ Level 3 (high strength)

Figure 76. Comparison of MSE of each model for three levels of D/t , d/t , and l^2/Dt

5.1.4 Proposed prediction models

As shown in Figure 75, the performance of the existing models changes with different levels of mechanical strength of the pipe material. This suggests that different failure pressure prediction models should be developed for different levels of mechanical strength. Figure 77(a) and Figure 77(b) display the scatter plots of σ_y vs. the grade, and σ_u vs. the grade in the dataset, respectively. Figure 77 shows that the correlations between σ_y or σ_u and the grade are similar: strongly positive correlated. This suggests that it is appropriate to develop the failure pressure prediction models based on the levels of either σ_y or σ_u , as both could be used to reflect the grade level. In this study, two sets of models are developed: one set of models is based on three levels of σ_y and another set based on three levels of σ_u . Then these two sets of models are compared to determine the final set.

For each level of σ_y or σ_u , the probabilistic failure pressure model follows the same multivariate linear regression formulation by adopting the existing prediction models as the independent variables:

$$Y = \theta_0 + \sum_{i=1}^m \theta_i \hat{y}_i + \sigma \varepsilon \quad (5.1)$$

where Y = predicted failure pressure or a suitable transformation; θ_i = model parameters; $\hat{y}_i$ = deterministic prediction from the 24 existing prediction models described in Appendix B; and $\sigma \varepsilon$ = model error in which σ is the standard deviation of the model error (assumed to be constant) and ε is the standard normal random variable (i.e., normality assumption). When considering all the existing prediction models in Eq. (5.1), the model is a full model. Since not all the terms contribute to the model prediction, a model selection procedure is adopted to eliminate the ones that do not contribute statistically significantly to the prediction.

Model development

With Eq. (5.1), the full model for each level of σ_y or σ_u is assessed based on the data within the corresponding level. In particular, randomly selected 80% of the data (or called training data) in each level is used for the model development, while the rest 20% of the data (or called test data) in each level is used for validation. With the full model, an all-possible subset model selection procedure is used to reduce the model size to determine the final formulation [34]. Since the defect width w is missing for 61% of the cases (i.e., 264 cases) in the database, the four existing models that use w are not considered in the model development. Hence, the full model has a model size of 20 (i.e., $m = 20$ in Eq. (5.1)), and the size of the reduced model varies from 1 to 19. In the all-

possible subset model selection, all possible combinations of predictors are evaluated for each model size (or subset) and the best model from that subset is identified. Then those best models from all the subsets are compared to determine the final model.

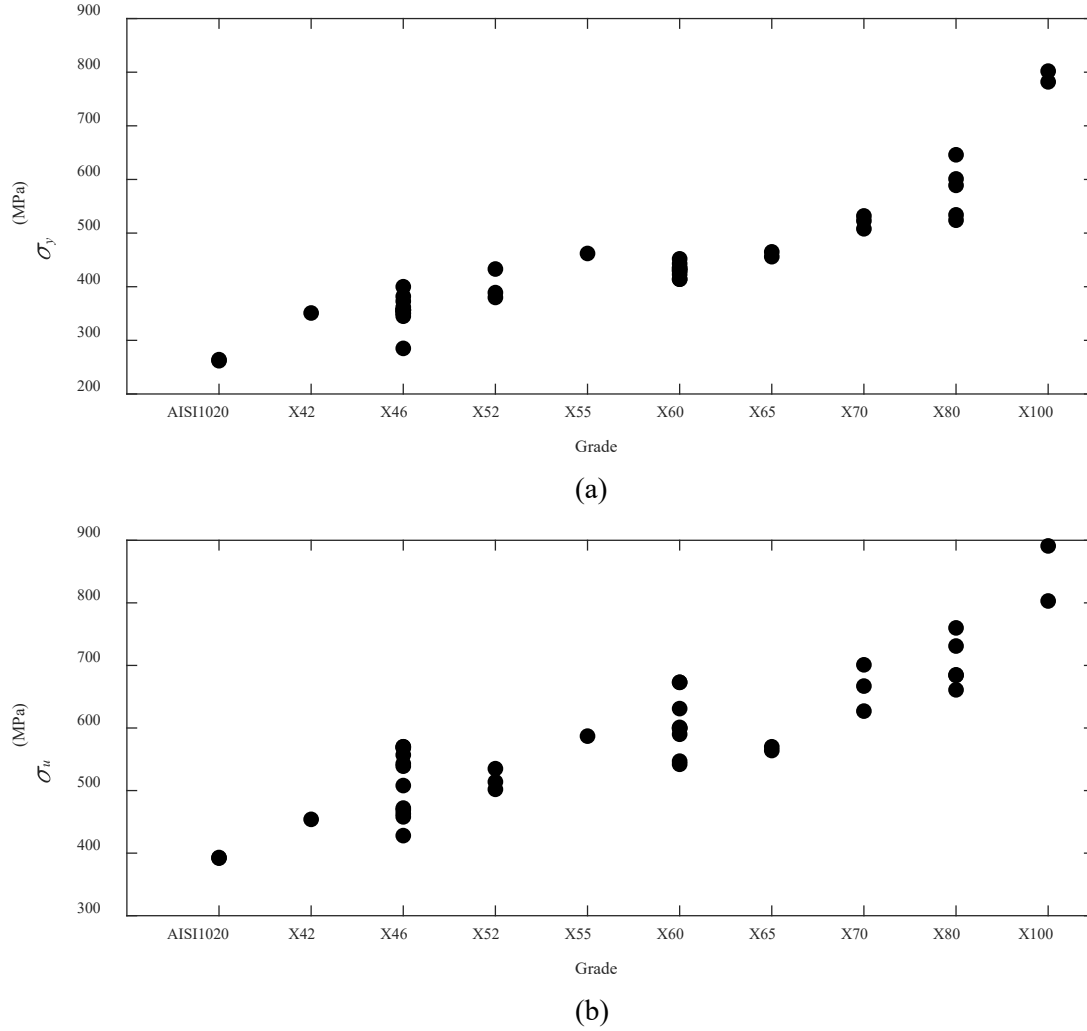


Figure 77. Scatter plots of (a) σ_y vs. grade and (b) σ_u vs. grade

To compare the model performance for each subset, this study uses three statistics measures: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and standard deviation of model error (σ). Both AIC and BIC measure how well the model fits the data through log likelihood, $\log(L)$, with the consideration of the number of predictors used in the model, trading off the complexity of the model formulation with its accuracy. These two measures are calculated as below:

$$AIC = 2k - 2 \log(L) \quad (5.2)$$

$$BIC = 2 \log(n)k - 2 \log(L) \quad (5.3)$$

where k = number of estimated parameters in the model, and n = number of data points. The less value of AIC or BIC is, the better performance the model has. The standard deviation of model error, σ , measures the dispersity in the model prediction; and the lower σ is, the better the model is.

For each subset (i.e., the possible models with the same model size), all the statistics measures advocate the same model as the best model. However, when comparing all the best models from all the subsets, these three statistics measures may suggest different models to be the most desirable one. Note that since the predictors, $\hat{y}_i$, in Eq. (5.1) are actually existing deterministic models that also have different extents of complexity in their formulations; thus, using just the number of predictors in the model as the measure of the model complexity in AIC and BIC may oversimplify the complexity measurement in this study. Consequently, the final model is determined by using engineering judgement for the complexity of the formula, σ for the accuracy in this study, and AIC and BIC as selection criteria references.

Probabilistic models

Instead of pre-defining the three levels of σ_y or σ_u , the ranges that define the levels are optimized by minimizing the total residuals of the three models of the set. For all the models, it is found that when the model size goes up to 3 or more, the change in σ is insignificant. This indicates that it is not beneficial to choose a model with a size larger than 2. Table 7 shows the model selection results for the two sets of developed models: one set based on three levels of σ_y , and the other set based on three levels of σ_u . For each level, Table 7 also shows the existing models selected resulted from the model selection and standard deviation of the model error, σ , for model sizes 1 and 2.

Table 7. Comparison of the best models of subsets

Level based on σ_y	Level 1 $262 \leq \sigma_y < 430$ MPa		Level 2 $430 \leq \sigma_y < 530$ MPa		Level 3 $530 \leq \sigma_y < 802$ MPa	
Model Size	1	2	1	2	1	2
Existing model selected	G6-24	G1-9, G3-15	G3-16	G1-7, G4-18	G3-16	G1-3, G1-4
σ (MPa)	2.0139	1.8781	1.7722	1.6195	1.2625	1.1025
Level based on σ_u	Level 1 $392 \leq \sigma_u < 600$ MPa		Level 2 $600 \leq \sigma_u < 700$ MPa		Level 3 $700 \leq \sigma_u < 891$ MPa	
Model Size	1	2	1	2	1	2
Existing model selected	G6-24	G4-18, G6-24	G1-5	G1-2, G4-18	G1-3	G1-3, G4-18
σ (MPa)	1.8442	1.8018	1.2253	1.0682	1.5616	1.3559

Table 7 shows that overall, the models developed based on the levels of σ_u have lower

model errors compared to the ones based on the levels of σ_y , except the models for Level 3 of σ_u . For model size 2 on Level 3, such difference is not too significant. Therefore, the models based on the three levels of σ_u are preferred. For each level, as expected, the model with model size 2 has a smaller σ value (i.e., more accurate) than the one with model size 1; however, when such decrease becomes marginal, the model with a smaller model size (i.e., less complexity in model formula) is preferred. Accordingly, the final selected models are highlighted with grey shown in Table 7. Table 8 shows the model formula and model parameter statistics of the final selected three models for the three levels of σ_u .

5.1.5 Model performance evaluation

Figure 78 provides the scatter plots of the point prediction of the proposed model, $\hat{Y}$ for each level of σ_u vs. the observed data, Y_{test} , that is the training data used for the model development. If the prediction is perfect, the dots should line up on the 1:1 line, shown as the solid line. The dashed lines are the mean ± 1 standard deviation of the model error. In these plots, the circle and cross markers refer to the data obtained from experimental and numerical burst tests, respectively. Figure 78 shows that the dots are evenly scatter around the 1:1 line for both the experimental and numerical data, indicating that the developed models provide unbiased predictions. The scatter degree of the dots reflects the accuracy of the model. For example, Figure 78(a) shows a slightly larger scatter compared to Figure 78(b) and Figure 78(c), indicating that the Level 1 model is less accurate than the Levels 2 and 3 models, consistent with the model errors shown in Table 8. Figure 78 also shows that most of experimental data is in Level 1, while only a few experimental data in Levels 2 and 3. Figure 78(a) indicates that the proposed model for Level 1 performs similarly for the experimental and numerical burst test data. However, as the experimental test data are very limited for Levels 2 and 3, one cannot conclusively evaluate the performance of the proposed models for the experimental cases for these two levels.

Figure 79 shows the performance of the proposed models compared with the existing models at the three levels of σ_u . As shown in Figure 79, regardless the levels, the proposed models (labeled as PM in Figure 79) are unbiased and have the lowest MSE. Note that Figure 79 indicates that the performances of the existing models at the optimized three levels of σ_u are similar to the ones at the three predefined levels of σ_y shown in Figure 75. The same model, G5-19, that was identified as the best existing model based on the levels of σ_y shown in Figure 75 is also the best

existing model based on the levels of σ_u shown in Figure 79. Interestingly, G5-19 is not selected in the proposed models through the model selection.

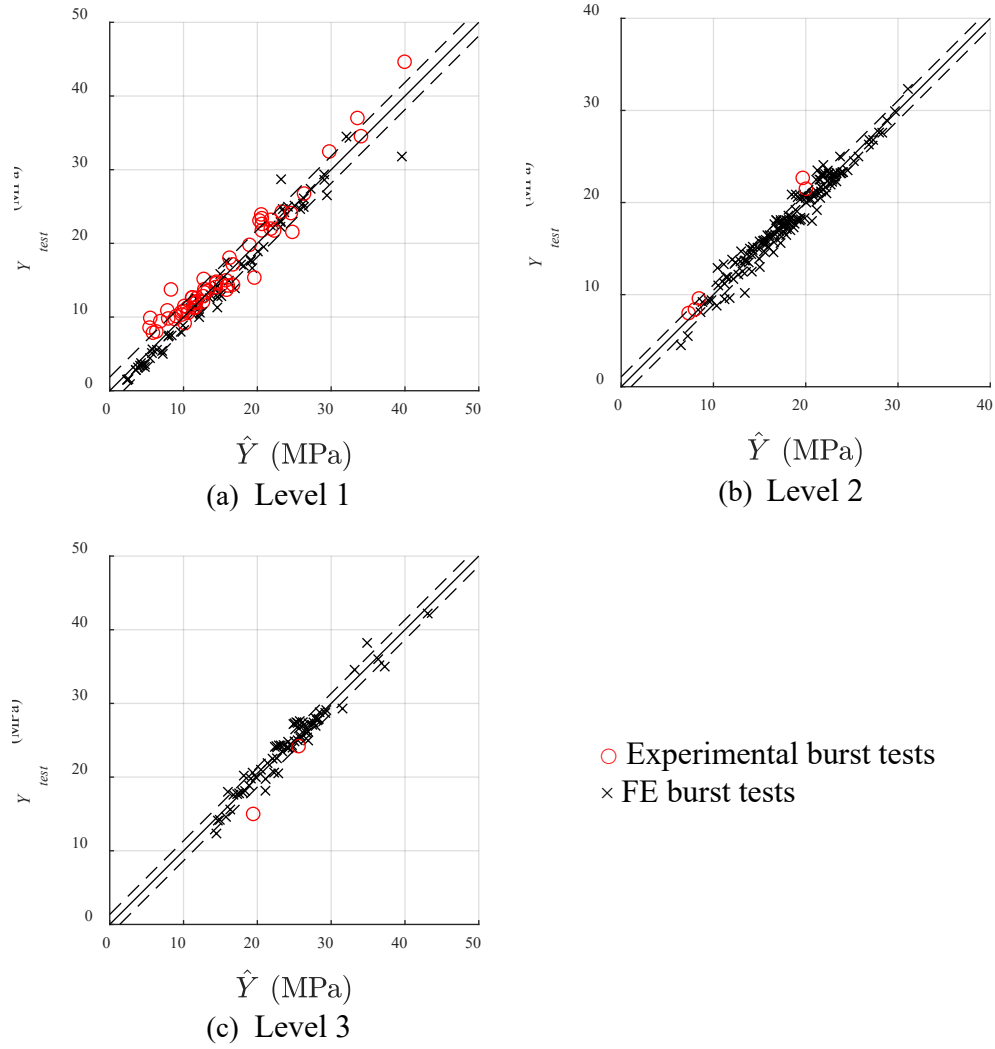


Figure 78. Scatter plots of the failure pressure predicted by the proposed models vs. the observed pressure

Table 8. Final selected model formula and model parameter statistics

Level based on σ_u	Formula	Model Parameters						
		θ_0		θ_1		θ_2		σ
		Mean	Std	Mean	Std	Mean	Std	
Level 1 $392 \leq \sigma_u < 600$ MPa	$\theta_0 + \theta_1 \hat{y}_{24}$	1.8469	0.3180	1.0281	0.0209	-	-	1.8442
Level 2 $600 \leq \sigma_u < 700$ MPa	$\theta_0 + \theta_1 \hat{y}_2 + \theta_2 \hat{y}_{18}$	-2.3322	0.3774	1.0751	0.0271	0.2978	0.0273	1.0682
Level 3 $700 \leq \sigma_u < 891$ MPa	$\theta_0 + \theta_1 \hat{y}_3 + \theta_2 \hat{y}_{18}$	3.4948	0.6490	0.9381	0.0501	0.2420	0.0518	1.3559

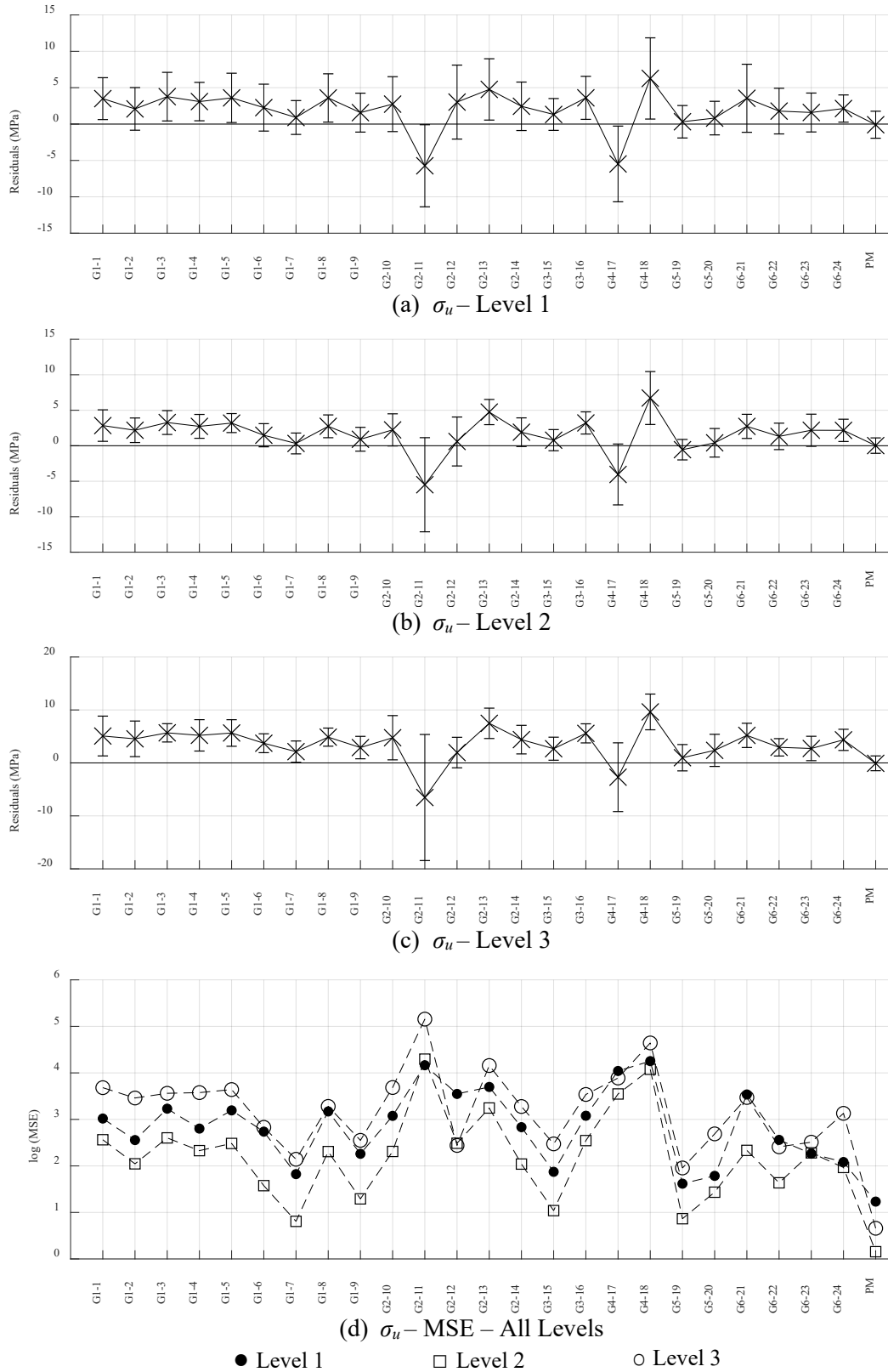


Figure 79. Comparison of residual and MSE of the proposed models with the existing models for three levels of σ_u

To further evaluate the proposed models' performance, Figure 80 shows the scatter plots of the point prediction of the proposed model or the besting existing model, G5-19, identified in Figure 79 vs. the observed data, Y_{test} using the validation data (i.e., the 20% of the data that is not used in the model development). Figure 80 displays that for each level the proposed model (marked as solid dots) has the unbiased predictions; and most of the solid dots are within the mean ± 1 standard deviation of the model error, which validate the proposed model. On the other hand, the scatter of the predictions from the best existing model (marked as crosses) shows only unbiased for Level 1, slightly overestimation for Level 2, and underestimation for Level 3. Furthermore, based on the scatter size, the prediction from G5-19 is very similar to the proposed model for Level 1, but the proposed model shows better accuracy for Level 2 and particularly Level 3. This indicates that the proposed models improve the accuracy from the existing models.

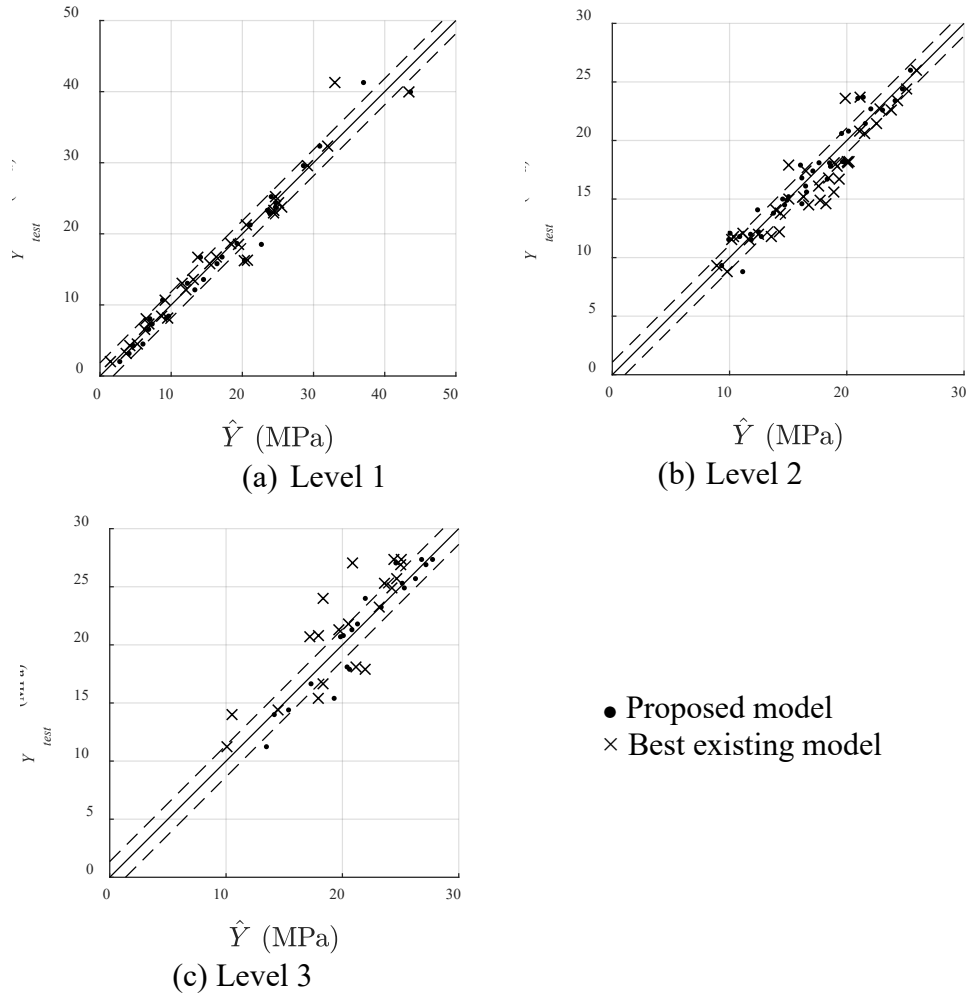


Figure 80. Scatter plots of the prediction of proposed model and best existing model vs. the validation data (20% of data)

5.1.6 Sensitivity analysis

Aforementioned, Figure 72 suggests that σ_y , D/t , d/t , l^2/Dt have effects on the failure pressure; therefore, a sensitivity analysis is performed to evaluate the impact of these quantities on the proposed model predictions. Since the proposed models are developed based on the levels of σ_u , only three quantities (i.e., D/t , d/t , l^2/Dt) are considered. For a comparison purpose, the sensitivities of these three quantities to the best existing model (i.e., G5-19) are also obtained.

Figure 81 shows the changes in the pressure prediction of a pipeline with $\sigma_u = 458$ MPa (belonging to Level 1), $D = 324$ mm, and $d = 2$ mm, when varying d/t , D/t , and l^2/Dt . For d/t and D/t (shown in Figure 81(a) and Figure 81(b) respectively), two cases are considered: one with a shorter defect length ($l = 100$ mm or $l^2/Dt = 5$) and the other one with a longer defect length ($l = 350$ mm or $l^2/Dt = 63$). For l^2/Dt (shown in Figure 81(c)), two cases are also considered: one with a shallow depth ($d = 2$ mm or $d/t = 33\%$) and the other with a deep defect depth ($d = 5$ mm or $d/t = 83\%$).

First, as expected, the predictions of both models decrease with the increase of the three quantities d/t , D/t , and l^2/Dt . This indicates that both the proposed and the best existing models are sensitive to the pipe thickness and the depth and length of the corrosion defect. The rates of burst pressure decrease for both models are overall very similar for all the cases except the sensitivity to l^2/Dt for the deep defect depth, where G5-19 has much greater rate of decrease than the proposed model when $l^2/Dt < 10$. For d/t and D/t , the sensitivity results are similar for both shorter and longer length defects. For l^2/Dt , the sensitivity depends on the depth of the defects: more sensitive for the deeper defect. In addition, for $l^2/Dt > 30$, the change rates in both model predictions become flatter, which means that the models become less sensitive for long defect lengths. Note that the sensitivity analysis for pipelines with σ_u in Levels 2 and 3 are found to be similar to the finding shown in Figure 81.

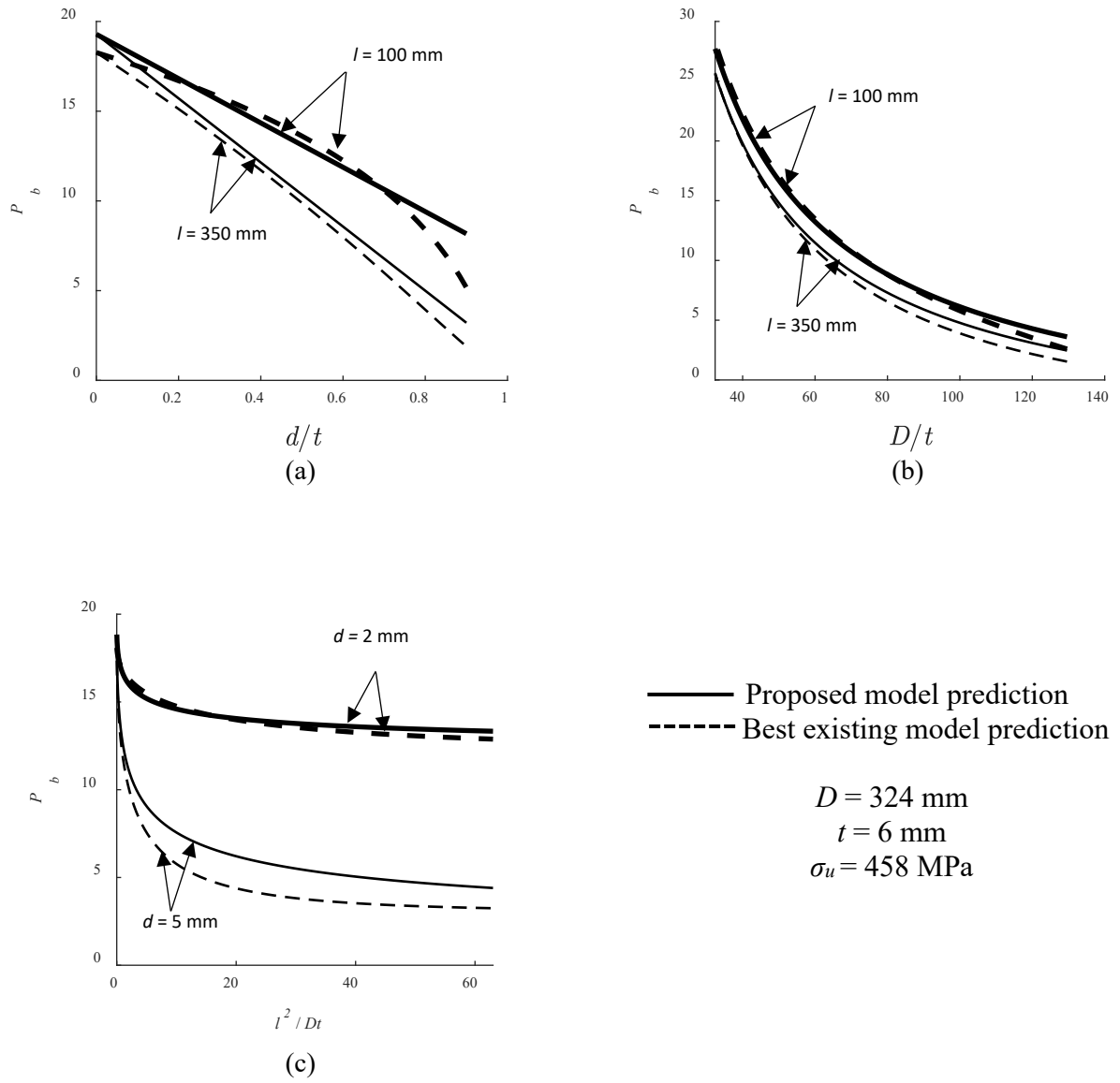


Figure 81. Sensitivity of the three selected quantities on pressure prediction

5.2 Probabilistic Models of Defect Interaction Identification and Failure Pressure for Pipelines with Interacted Corrosion Defects

5.2.1 Background

Many interaction rules have been developed in the past to identify defect interaction of adjacent defects, such as Kiefner and Vieth (KV) [26], POF [27], DNV RP F101 [8], API RP 579 [28], BS 7910 [29], 6WT [30] and 3WT [5]. Recently, Li et al. [31] proposed an interaction rule expressed in terms of pipe diameter and/or thickness based on finite element (FE) analysis of models with short and long corrosion defects. Their interaction rule provides different formulations for different ranges of corrosion defect length; but the FE models only contain the adjacent defects with the same length, meaning the developed formula may not be suitable when the adjacent defects have different lengths. Mondal and Dhar [32] proposed an interaction rule for longitudinal spacing; however, the rule is developed using the adjacent defects with the same depth, meaning the rule may not be appropriate when the adjacent defects have different depths. In addition, Zhang and Tian [33] developed an interaction rule for longitudinal spacing considering defect depth and steel grade, but their interaction rule has the same limitation as the one proposed by Mondal and Dhar [32].

To evaluate the performance of the existing rules, several studies have been conducted. For instance, Benjamin et al. [34] compared the performance of KV [26], POF [27], and DNV RP F101 [8] based on 26 finite element models of short corrosion defects. They found that the DNV interaction rule had the best performance. Similarly, Li et al. [31] compared their proposed interaction rule with 5 interactions rules (i.e., KV [26], POF [27], DNV RP F101 [8], API RP 579 [28], and BS 7910 [29]), and the results show that their proposed interaction rules performed the best. In summary, although many interaction rules have been developed to identify defect interaction, the existing interaction rules are not consistent; and in addition, all of them are typically deterministic, which are not able to capture the inherent uncertainty in the defect interaction. Therefore, a probabilistic model that considers all the influencing factors (e.g., the defect depth, length and width, and pipeline strength, thickness, and diameter) is needed.

Regarding failure pressure prediction, many models have been developed for pipelines with single corrosion defect, and only a few models were developed considering interacting defects. Nevertheless, the models developed for single corrosion defect such as the ASME B31G method [5], the RSTRENG 085dL method [5], and the DNV RP-F101 method for single defects

(Part B) [8] have been adopted to calculate the failure pressure of pipelines with interacting defects by using the maximum depth and overall length of the colony in those models. Moreover, model such as the RSTRENG Effective Area [5] developed for isolated complex-shaped defect can be used to calculate the failure pressure of colony of defects by incorporating the length of full wall-thickness pipe that separates the adjacent defects as a part of the depth profile of the complex-shaped defect [35]. Det Norske Veritas (DNV) with the cooperation of the BG Technology developed a method called DNV RP-F101 for interacting defects (Part B) [8]. In this method, the failure pressure of the colony corrosion defects is determined by the minimum value of all pressures based on each single defect within the colony and each possible combination of the adjacent defects. Note that the DNV approach cannot be implemented to the cases where individual profiles overlap when projected onto the longitudinal plane [36]. Benjamin et al. [35] proposed the Mixed-Type Interaction (MTI) method based on the DNV RP-F101 method for interacting defects (Part B) by considering the length of full wall-thickness pipe that exists between each pair of defects within a colony of random configuration; and it also considers the interacting defects whose individual profiles overlap when projected onto the longitudinal plane. In addition, Chen et al. [37] developed a method to predict the failure pressure for high strength pipeline with multiple corrosion defects, which follows the same procedure described above for the DNV RP-F101 method for interacting defects (Part B) but with a different formulation to calculate the failure pressure for single defect and effective depth of the combined defects.

The performance of some of the existing prediction models has also been compared by researchers in the literature. For example, Benjamin et al. [35] used laboratory tests to compare five assessment methods, the ASME B31G method [5], the DNV RP-F101 method for single defects (Part B) [8], the RSTRENG Effective Area method [5], the DNV RP-F101 method for interacting defects (Part B) [8], and the MTI method [35]; and the MTI method provides the most accurate predictions. In general, most of the models are found to be conservative, indicating when they are used in the pipeline risk management, they will lead to unnecessary repairs and maintenance.

5.2.2 Existing model formulations

Existing interaction models

An interaction rule is a rule used to define if two adjacent defects interact and is generally expressed in terms of the spacing between the two individual defects (either longitudinal spacing

or circumferential spacing). Figure 82 shows a configuration of two defects (Defect 1 and Defect 2) with the relevant geometric quantities. Most of the existing interaction rules can be written as [38]:

$$S_L \leq (S_L)_{Lim} \text{ and } S_C \leq (S_C)_{Lim} \quad (5.4)$$

where S_L = longitudinal spacing, S_C = circumferential spacing, $(S_L)_{Lim}$ and $(S_C)_{Lim}$ = longitudinal spacing limit and circumferential spacing limit defined by the interaction rule, respectively. If there are more than two defects in the colony, the interaction rule is then applied for all possible pairs of adjacent defects [36]. Table 9 provides the formulations for calculating $(S_L)_{Lim}$ and $(S_C)_{Lim}$ by various existing interaction rules in the literature. As shown in Table 9, some use the defect depth, length, or width in the spacing limits (e.g., POF [27]), while others use just pipeline thickness and/or diameter (e.g., KV [26]). The variation in the existing formulations shown in Table 9 indicates that a consistent rule is still needed, and all the influencing factors (including the defect geometries and pipeline properties) should also be examined holistically. In addition, the existing interaction rules are deterministic, which are not able to capture the inherent uncertainty in the defect interaction.

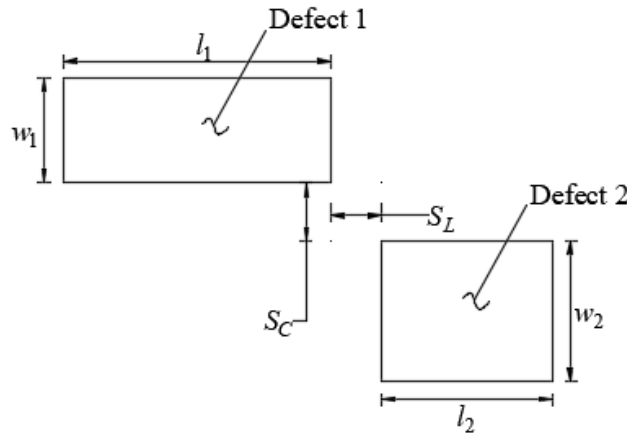


Figure 82. Configuration of two defects

Existing capacity pressure prediction models

The failure pressure of a pipe with a colony of corrosion defects is typically evaluated using an interaction rule and a Level-1 assessment method [35]. A Level-1 assessment method uses the maximum depth and projected axial length of a defect to evaluate the remaining strength (i.e., burst pressure capacity, P_b) [39]. When there is no defect interaction, the pressure capacity is determined by the smallest value of all pressures based on each single defect within the colony. In the

literature, the most commonly used Level-1 assessment methods are the ASME B31G method [5], the RSTRENG 085dL method [5] , and the DNV RP-F101 method for single defects (Part B) [8]; the formulations of these three prediction models are summarized in Table 10, labeled as L1-1, L1-2, and L1-3, respectively.

Table 9. Existing interaction rules

Interaction rules	$(S_L)_{Lim}$	$(S_C)_{Lim}$
KV [26]	25.4 mm	$6t$
POF [27]	$\min(6t, l_1, l_2)$	$\min(6t, w_1, w_2)$
DNV RP F101 [8]	$2\sqrt{Dt}$	$\pi\sqrt{Dt}$
API RP 579 [28]	$(l_1 + l_2) / 2^*$	$(w_1 + w_2) / 2^*$
BS 7910 [6]	$2\sqrt{Dt}^*$	$3\sqrt{Dt}^*$
6WT [30]	$6t$	$6t$
3WT [5]	$3t$	$3t$
Mondal and Dhar [32]	$(15.91 - 7.69d/t)t$ or $(3 - 1.46d/t)\sqrt{Dt}$	-
Zhang and Tian [33]	$(3954.875 - 3956.144 \cdot 0.994^{d/t})t$ for X65 or $(27.831 - 33.22 \cdot 0.132^{d/t})t$ for X80	-
Li et al.[31]	$2\sqrt{Dt}$, $l/\sqrt{Dt} \leq \sqrt{20}$ $\sqrt{Dt}$, $\sqrt{20} < l/\sqrt{Dt} \leq \sqrt{50}$ t , $l/\sqrt{Dt} > \sqrt{50}$	$0.1\pi D$, $l/\sqrt{Dt} \leq \sqrt{20}$ $0.05\pi D$, $\sqrt{20} < l/\sqrt{Dt} \leq \sqrt{50}$ t , $l/\sqrt{Dt} > \sqrt{50}$
t = thickness of pipeline D = outside diameter of pipeline d = identical defect depth for the two defects l = identical defect length for the two defects l_1, l_2 = defect lengths for the two defects, respectively w_1, w_2 = defect widths for the two defects, respectively		
*This formulation is an interpretation from the original reference by reference [31].		

Table 10. Existing capacity pressure prediction models

Method	Formulation
ASME B31G (L1-1)	$P_b = \begin{cases} \frac{2t}{D} (1.1\sigma_y) \left[\frac{1 - (2/3)(d/t)}{1 - (2/3)(d/t)M^{-1}} \right], & l^2/Dt \leq 20 \\ \frac{2t}{D} (1.1\sigma_y) [1 - (d/t)], & l^2/Dt > 20 \end{cases} \quad (5.5)$ $M = \sqrt{1 + 0.8 \left(\frac{l^2}{Dt} \right)} \quad (5.5a)$
RSTRENG 085dL (L1-2)	$P_b = \frac{2t}{D} (\sigma_y + 69[MPa]) \left[\frac{1 - 0.85(d/t)}{1 - 0.85(d/t)M^{-1}} \right] \quad (5.6)$ $M = \begin{cases} \sqrt{1 + 0.6275 \left(\frac{l^2}{Dt} \right) - 0.003375 \left(\frac{l^2}{Dt} \right)^2}, & l^2/Dt \leq 50 \\ 3.3 + 0.032 \left(\frac{l^2}{Dt} \right), & l^2/Dt > 50 \end{cases} \quad (5.6a)$
DNV RP-F101 for single defects (Part B) (L1-3)	$P_b = \frac{2t\sigma_u(1 - (d/t))}{(D - t) \left(1 - \frac{(d/t)}{M} \right)} \quad (5.7)$ $M = \sqrt{1 + 0.31 \left(\frac{l^2}{Dt} \right)} \quad (5.7a)$
σ_y = yield strength of pipe material, σ_u = ultimate strength of pipe material, d : maximum depth of the corrosion defect, M = folias or bulging factor	

Table 11. Summarized procedure for DNV RP-F101 for interacting defects (Part B) and MTI method

Calculate the failure pressure of each single defect within the colony of defects, P_i	$P_i = \frac{2t\sigma_u(1 - (d_i/t))}{(D - t)\left(1 - \frac{(d_i/t)}{M_i}\right)} \quad i: 1 \dots N \text{ defects} \quad (5.8)$ $M_i = \sqrt{1 + 0.31\left(\frac{l_i^2}{Dt}\right)} \quad (5.8a)$
Calculate the combined length of all combinations of adjacent defects, l_{jk}	$l_{jk} = l_k + \sum_{i=j}^{i=k-1} (l_i + S_{Li}) \quad j, k = 1 \dots N \quad (5.9)$ where S_{Li} = longitudinal spacing between adjacent defects
Calculate the effective depth of all combinations of adjacent defects, d_{jk}	DNV RP-F101 for interacting defects (Part B) $d_{jk} = \frac{\sum_{i=j}^{i=k} d_i l_i}{l_{jk}} \quad (5.10)$
	MTI method $d_{jk} = \frac{V_{jk}}{l_{jk} w_{jk}} \quad (5.11)$ $V_{jk} = \sum_{i=j}^{i=k} d_i l_i w_i \quad j, k = 1 \dots N \quad (5.11a)$ $w_{jk} = w_k + \sum_{i=j}^{i=k-1} (w_i + S_{Ci}) \quad j, k = 1 \dots N \quad (5.11b)$ where S_{Ci} = circumferential spacing between adjacent defects
Calculate the failure pressure of all combinations of adjacent defects, P_{jk}	$P_{jk} = \frac{2t\sigma_u(1 - (d_{jk}/t))}{(D - t)\left(1 - \frac{(d_{jk}/t)}{M_{jk}}\right)} \quad (5.12)$ $M_{jk} = \sqrt{1 + 0.31\left(\frac{l_{jk}^2}{Dt}\right)} \quad (5.12a)$
Determine burst pressure, P_b	$P_b = \min(\forall P_i, \forall P_{jk}) \quad (5.13)$

Alternatively, Level-2 assessment methods can be used for the failure pressure prediction for the pipe with interacting corrosion defects. A Level-2 assessment method evaluates the remaining strength by considering the possibility of interaction effect among defects and/or the impact of the defect shape [39]. Some of the Level-2 assessment methods are the RSTRENG

Effective Area method [5] (labeled as L2-4), the DNV RP-F101 method for interacting defects (Part B) (labeled as L2-5) [8], and the MTI method [35] (labeled as L2-6). The RSTRENG Effective Area method [5] developed by the ASME involves numerous measurements of the depth of the corrosion all over the corroded area, requiring therefore a longitudinal profile of the corroded area [5]; and the pressure prediction is calculated by:

$$P_b = \frac{2t}{D}(\sigma_y + 69[MPa]) \left[\frac{1 - A/A_0}{1 - (A/A_0)M^{-1}} \right] \quad (5.14)$$

$$M = \begin{cases} \sqrt{1 + 0.6275 \left(\frac{l^2}{Dt} \right) - 0.003375 \left(\frac{l^2}{Dt} \right)^2}, & l^2/Dt \leq 50 \\ 3.3 + 0.032 \left(\frac{l^2}{Dt} \right), & l^2/Dt > 50 \end{cases} \quad (5.14a)$$

where A : longitudinal area of metal loss and A_0 : original uncorroded area of length l and thickness t . The failure pressure is obtained by evaluating all possible combinations of local metal loss with respect to original material using iteration and the lowest calculated failure pressure is retained as the failure pressure of the colony corrosion defects [5].

Compared to the RSTRENG Effective Area method, the other two existing methods involve several steps, which are summarized in Table 11, where the failure pressure is determined by the minimum value of all pressures based on each single defect within the colony and all possible combinations of the adjacent defects. Figure 83 adapted from DNV [8] shows a combined Defect jk defined by single Defect j to single Defect k with the defect depth profiles projected on the longitudinal plane. The only difference between these two methods is how the effective depth d_{jk} is calculated, where the DNV RP-F101 method for interacting defects (Part B) uses the surface area and the MTI uses the volume of the metal loss of individual defect within the combined defects. Lastly, using the minimum value of all calculated pressures also implies that these two existing models suggest the existence of the defect interaction when one of P_{jk} is the lowest calculated failure pressure.

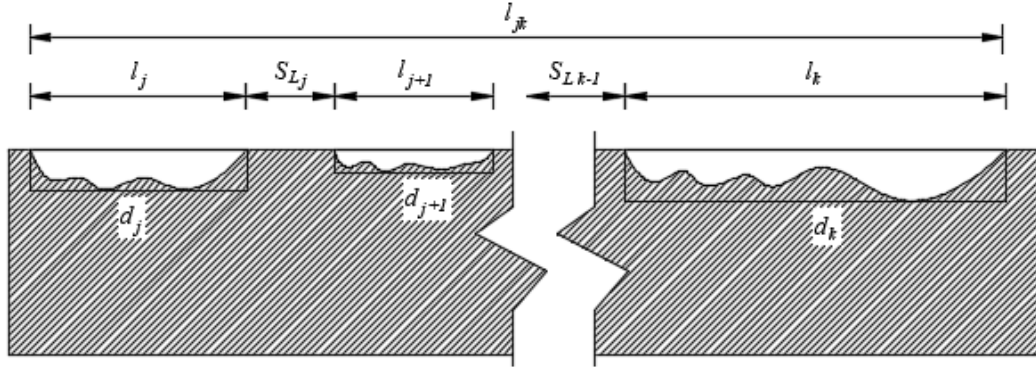


Figure 83. Combining interacting defects

5.2.3 Data Collection

A comprehensive failure pressure data is established in this study, which is then used to evaluate the performance of the existing interaction rules and prediction models, and also to develop accurate interaction rule and failure pressure assessment method for pipeline with a colony of corrosion defects. The database established consists of the data directly collected from literature and additional numerical data obtained from FE analysis conducted in this study.

Data collection from literature

A total of 202 different burst test results are collected from past studies, out of which 25 are from laboratory experimental burst test and 117 are from FEM simulations. Table 12 provides a summary of the data collected, where $d_{cluster}$, $l_{cluster}$, and $w_{cluster}$ are the depth (maximum), length, and width of a cluster of defects, respectively. Since the failure pressure of a pipe with a colony of defects depends on the pipe material and geometry properties, defect geometry, and defect spacing, it is worthy to examine the correlation of the following quantities with the failure pressure (P_b) through scatter plots shown in Figure 84 using the data collected: yield strength (σ_y), ratio of pipe diameter to pipe thickness (D/t), ratio of maximum defect depth within a cluster of defects to pipe thickness (d_{max}/t), ratio of cluster defect depth to pipe thickness (d_{vol}/t) where d_{vol} is calculated from the volume of metal loss of the cluster (i.e., $d_{vol} = V_{clus}/(w_{clus} \cdot l_{clus})$), ratio of cluster length squared to the multiplication of pipe diameter and thickness (l_{clus}^2/Dt), product of cluster width and cluster length ($w_{clus} \cdot l_{clus}$), and ratio of cluster width to pipe diameter (w_{clus}/D).

In Figure 84, the circle and cross markers refer to the data obtained from experimental and numerical burst tests, respectively. Figure 84(a), the scatter plot of σ_y vs. P_b , shows that an overall increase of σ_y leads to a higher burst pressure, except a few cases circled by the dotted lines. These

cases may be explained by their low values of D/t (referring to very thick-wall pipes) as shown in the dotted circle in Figure 84(b). Figure 84(b) on the other hand indicates that a low D/t ratio (referring to very thick-wall pipes) leads to a high burst pressure, as expected. Figure 84(c), (d), (f) and (g) show the scatter plots of d_{max}/t , d_{vol}/t , $w_{clus} \cdot l_{clus}$, and w_{clus}/D over P_b , respectively; no distinct trends between P_b and these quantities are found. However, Figure 84(e) indicates a negative correlation between P_b and $\log(l_{clus}^2/Dt)$.

Table 12. Summary of the database collected from literature

Grade	D/t	σ_y (MPa)	σ_u (MPa)	$d_{cluster}/t$	$l_{cluster}$ (mm)	$w_{cluster}$ (mm)	Counts
X60	[30.00 - 57.08]	[435 -452]	[542 - 560]	[0.25 - 0.80]	[35.00 - 475.60]	[35.00 - 418.39]	81
X70	[57.25 - 58.05]	[580 - 662]	[728 - 773]	[0.58 - 0.63]	[110.00 – 430.00]	[60.00 – 170.00]	16
X80	[56.64 – 57.43]	[534 - 589]	[661- 731]	[0.30 - 0.69]	[39.60 - 1072.20]	[31.90 - 383.90]	105

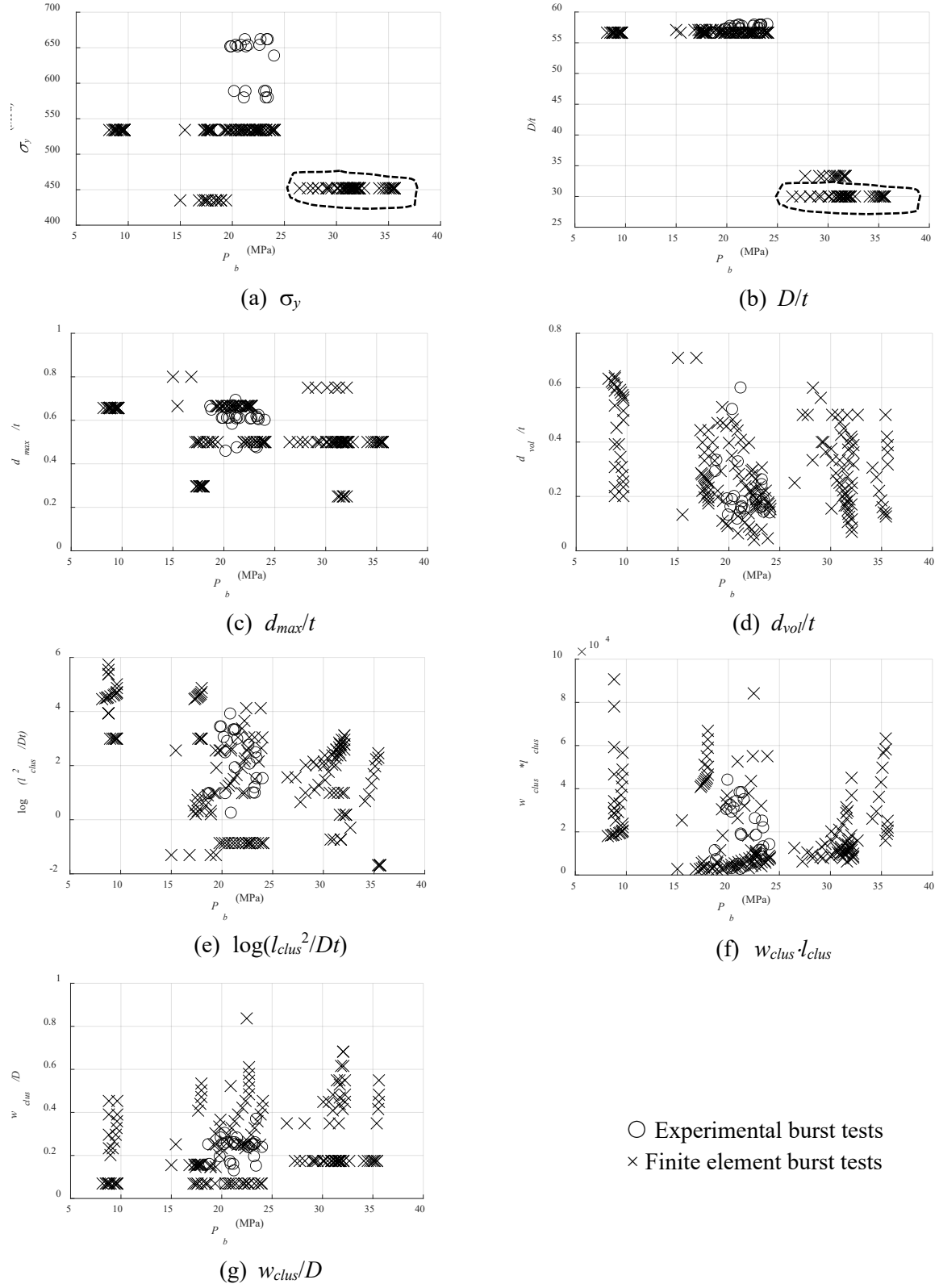


Figure 84. Scatter plots of burst pressure (P_b) vs. selected quantities

Additional numerical data

As shown in Table 12, the existing data are only for three grades (i.e., X60, X70, and X80) and not for low grades (e.g., X42, X46, X52, or X55). To complement the existing data collected in the literature, FEMs are used to generate additional cases. In this study, the FEMs are developed in ABAQUS. For computational efficiency, the corrosion defect is modeled as a rectangular shape as shown in Figure 85. ABAQUS Statics-General procedure is used for the analysis, and the burst pressure is determined when the von Mises stress at any point of the defect area reaches the ultimate tensile strength of pipe steel [22].

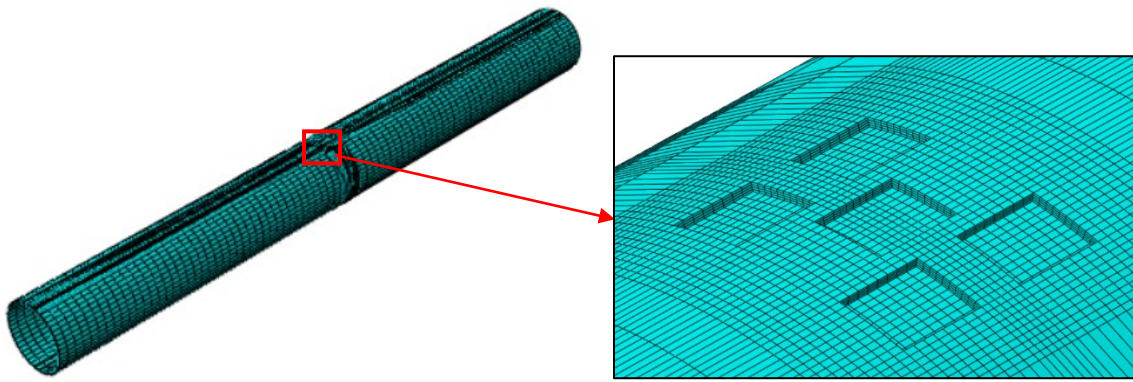
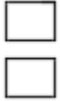
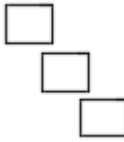
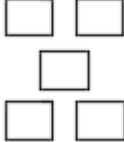


Figure 85. A corroded pipe modeled in ABAQUS

A few laboratory test results from the literature are selected for the FEM validation. Table 13 shows the summary of selected test cases and the burst pressure comparison of the test results and the FEM simulation results. Table 13 shows the defects spacing configuration of the selected test cases, and also compares the failure pressure data reported in the literature ($P_{b,test}$) and the failure pressure obtained from the FE analysis ($P_{b,FE}$) conducted in this study. In Table 13, all the selected test cases are grade X80 from [40] and [41], in which the stress-strain curves required to run the FEM analysis are provided.

Table 13. Selected cases for FEM validation

Grade	Specimen	D (mm)	Group of defects configuration	t (mm)	σ_y (MPa)	σ_u (MPa)	$(P_{b,test})$ (MPa)	$(P_{b,FE})$ (MPa)	Error (%)
X80	IDTS 3 [19]	458.8		8.10	534.1	661.4	20.31	20.48	-0.8223
X80	IDTS 4 [19]	458.8		8.10	534.1	661.4	21.14	22.61	-6.9328
X80	IDTS 9 [20]	459.4		8.00	589.0	731.0	23.06	22.97	0.4076
X80	IDTS 10 [20]	459.4		8.00	589.0	731.0	23.23	22.97	1.1365

As shown in Table 13, the error percentages between $P_{b,test}$ and $P_{b,FE}$ for all cases are within 10%. Therefore, the FE models are validated and can be used to assess the failure pressure for other defect scenarios. To cover a wide range of grade in the numerical cases, the stress-strain curves provided in the collected database with isolated defects are used to generate the FE models with colony defects, since the material stress-strain curves are independent of defects. With the validated FE models, a total of 783 new numerical cases are generated. These added numerical cases are designed to cover a wider range of yield strength (σ_y ranging from 262 MPa to 782 MPa) and four quantities: D/t , d_{max}/t , l_{clus}^2/Dt , and w_{clus}/D . Figure 86 shows the scatter plots of σ_y vs the four quantities based on the experimental and FE burst test data collected from literature and the new FEM data. As shown in Figure 86, the new cases are designed to cover the regions where the data collected from literature is scarce. It is worth stating that most of the total number of generated numerical cases (783) is originally generated in this study to understand the impact of various quantities (e.g., spacing, defect sizes, material properties) on the defect interaction. Table 14 provides a summary of the numerical data generated in this study and the ranges covered in terms of D/t , σ_y , σ_u , and depth (maximum), length, and width of the cluster of defects.

There are three situations in the database: (1) cases with interacted defects, (2) cases without interacted defects, and (3) cases in which defect interaction cannot be determined. In this

study, adjacent defects are considered to be interacted when the ratio of the actual burst pressure, P_a , to the burst pressure due to isolated defects, P_i , is less than 0.99 [31]. Table 15 summaries the numbers of the data points for each case in the total database (including the data collected from literature and the numerical data generated from this study). As shown in Table 15, most of the data are the cases with interacted defects; and also the colonies with two defects are the majority.

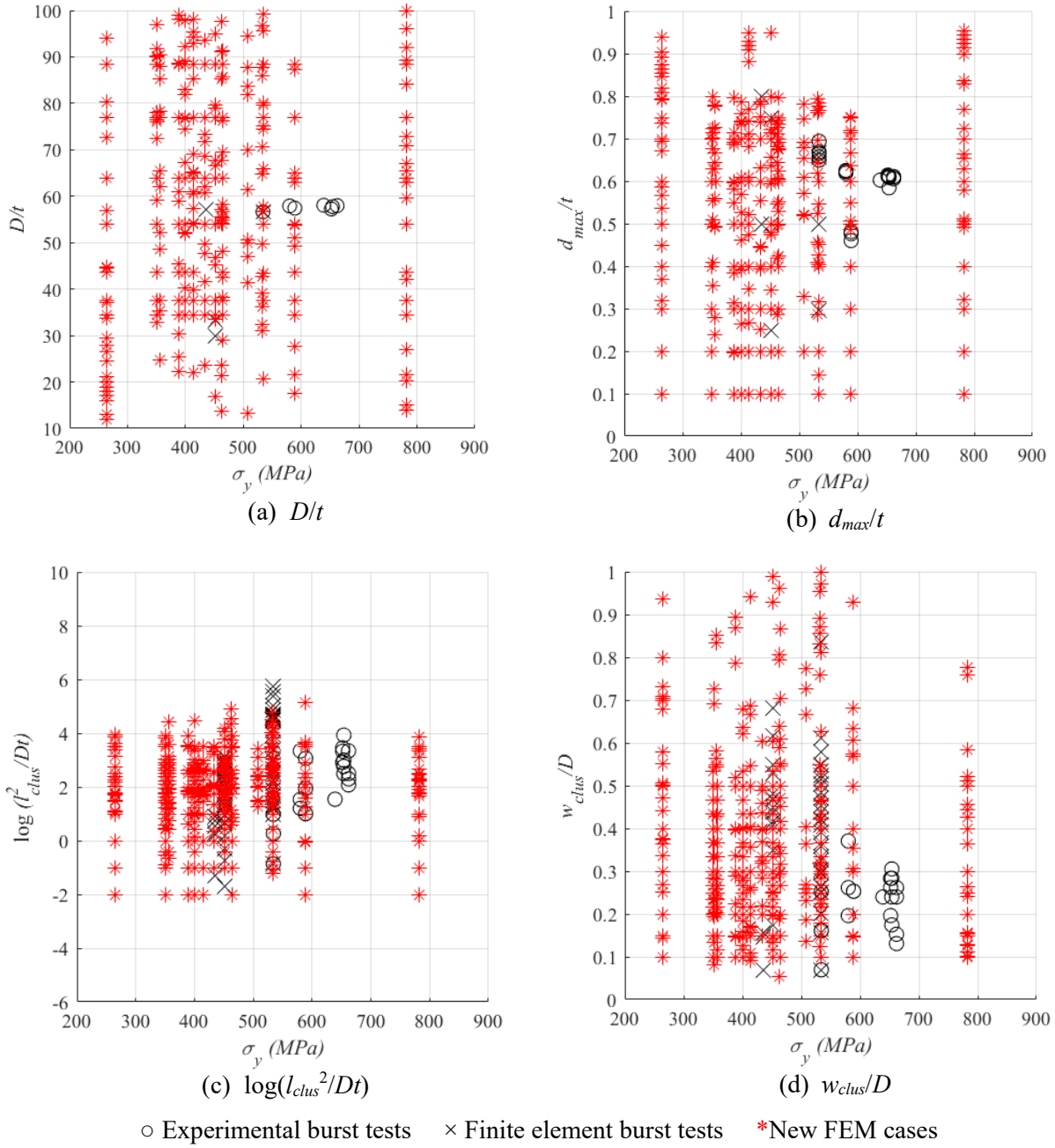


Figure 86. Scatter plots of selected quantities vs. yield stress (σ_y)

Table 14. Summary of the database generated from FE

Grade	D/t	σ_y (MPa)	σ_u (MPa)	$d_{cluster}/t$	$l_{cluster}$ (mm)	$w_{cluster}$ (mm)	Counts
AISI 1020 Mild	[12.00 - 94.00]	264	392	[0.10 - 0.94]	[42.48 - 489.13]	[61.00 - 320.01]	29
X42	[32.87 - 96.98]	351	454	[0.10 - 0.80]	[40.00 - 517.37]	[30.00 - 426.05]	115
X46	[24.65 - 97.95]	[356 - 400]	[469 - 508]	[0.10 - 0.79]	[40.00 - 501.90]	[30.00 - 497.67]	227
X52	[22.23 - 98.97]	[389 - 433]	[502 - 535]	[0.10 - 0.80]	[40.48 - 462.10]	[76.20 - 359.03]	34
X55	[13.76 - 97.57]	462	587	[0.29 - 0.73]	[82.66 - 468.59]	[74.37 - 488.58]	10
X60	[16.97 - 98.00]	[414 - 452]	[542 - 600]	[0.10 - 0.95]	[40.00 - 439.09]	[30.00 - 421.38]	171
X65	[28.98 - 91.45]	465	564	[0.10 - 0.80]	[42.48 - 425.67]	[76.20 - 441.23]	19
X70	[13.23 - 96.77]	[508 - 532]	[627 - 667]	[0.20 - 0.79]	[107.13 - 585.86]	[95.52 - 422.29]	23
X80	[17.60 - 99.31]	[534 - 589]	[661 - 731]	[0.10 - 0.78]	[40.00 - 493.24]	[30.00 - 399.31]	129
X100	[14.00 - 100.00]	782	803	[0.10 - 0.96]	[42.00 - 474.44]	[47.00 - 481.78]	26

Table 15. Number of data points

Number of defects in the colony	With interaction	Without interaction	Indeterminate interaction
2 defects	453	353	20
3 defects	111	0	4
4 defects	16	1	2
More than 4 defects	10	0	15

5.2.4 Performance comparison of existing models

The established database of pipelines is used to evaluate the performance of the existing interactions rules and failure pressure prediction models for colony of corrosion defects. Only the data with two defects are used to compare the performance of existing interaction rules, while all the data are used to compare the performance of existing failure pressure models.

Existing interaction models

The performance of an interaction rule can be quantified using a hit or miss approach. In the hit or miss approach, there are four possible outcomes: true positive (TP) - identifying no interaction when there is no interaction, true negative (TN) - identifying interaction when there is interaction,

false positive (FP) - identifying no interaction when there is an interaction, and false negative (FN) - identifying interaction when there is no interaction. Then the probability of correct detection, P_{CD} , is used as model prediction accuracy and is calculated as:

$$P_{CD} = \frac{n_{TP} + n_{TN}}{n_{total}} \quad (5.15)$$

where n_{TP} = number of TP tests, n_{TN} = number of TN tests, and n_{total} = total number of tests. Apparently, a good model should have a high value of P_{CD} .

Figure 87 displays the performance of seven existing interaction rules: KV [26], POF [27], DNV RP F101 [8], API RP 579 [31], BS 7910 [31], 6WT [30], and 3WT [5]. Note that not all the existing interaction rules listed in Table 9 are compared in Figure 87 due to their inapplicability to most of the data used in this study. For example, the interaction rules purposed by Mondal and Dhar [32], and Zhang and Tian [33] are only suitable to the cases where the adjacent defects have the same depth. As shown on Figure 87, POF, API RP 579, and 6WT interaction rules have better performance compared to the rest of the interaction rules. Both model POF and API RP 579 are based on defects length and width, while the interaction rule 6WT is based on only pipe thickness, t . Even though these three interaction rules have similar performance, there is a lack of consistency in the formulations of the spacing limits.

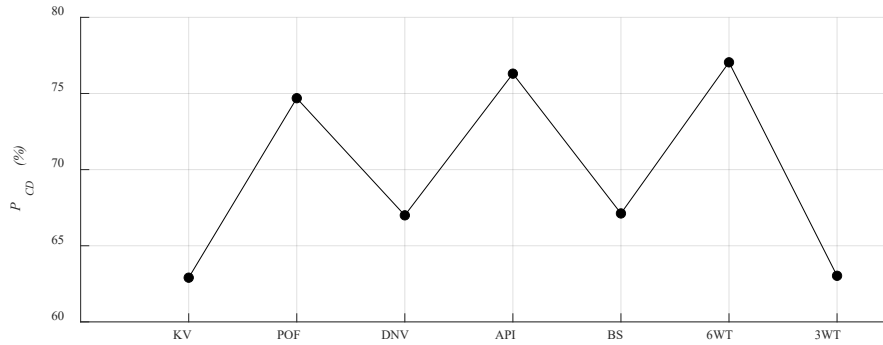


Figure 87. Comparison of the existing interaction rules using P_{CD}

Existing prediction models

The performance comparison of the existing failure prediction models for pipeline with colony of defects is conducted using the mean, standard deviation of the ratio of the predicted to the actual burst failure pressures, P_b/P_a . With the interaction rule 6WT, three Level-1 assessment methods (i.e., the ASME B31G method [5], the RSTRENG 085dL method [5], and the DNV RP-F101 method for single defects (Part B) [8]), labeled as L1-1, L1-2, and L1-3, respectively are compared

with three Level-2 assessment methods (i.e., the DNV RP-F101 method for interacting defects (Part B) [8], the RSTRENG Effective Area method [5], and the MTI method [35]), labeled as L2-4, L2-5, and L2-6, respectively. Since the DNV RP-F101 method is only applicable to the cases in which the individual defect profiles do not overlap projecting onto the longitudinal plane, the model performance comparison is done using two data sets: the 1st set refers to the subset of 698 data where the individual defect profiles do not overlap projecting onto the longitudinal plane and the 2nd data set refers to all the data (i.e., 985 data). The comparison results are shown in Figure 88, where the crosses refer to mean of P_b/P_a and the horizontal bars refer to mean ± 1 standard deviation.

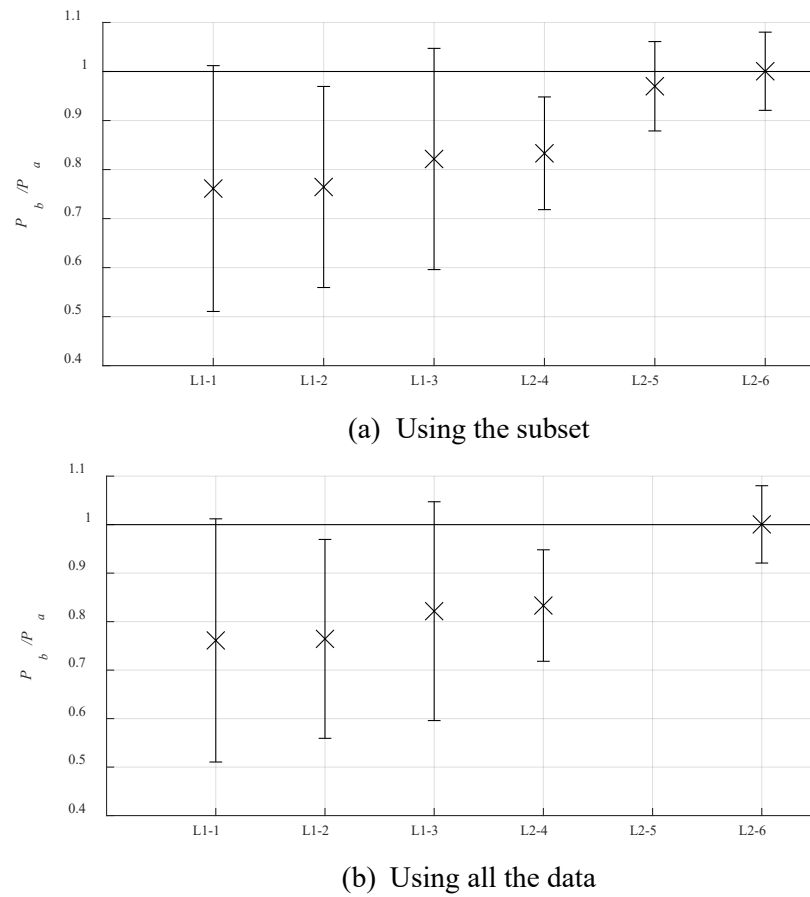


Figure 88. Comparison of existing failure pressure models

As shown in Figure 88(a), all the 6 models except L2-6 underestimate the failure pressure, resulting in mean of P_b/P_a lower than 1. Also, Figure 88(a) clearly indicates that the Level-2 assessment methods perform better than the Level-1 assessment methods. This shows that

assessment methods that use defect depth profile and consider all combinations of adjacent defects give better failure prediction than using interaction rules. In addition, the MTI method (L2-6) performs the best for the 1st data set. When using all the data, Figure 88(b) compares the performance of 5 models and the DNV RP-F101 method (L2-5) is not included in the comparison. Similar to the results shown in Figure 88(a), all the models except L2-6 underestimate the failure pressure; the Level-2 assessment methods have better performance; and the MTI method (L2-6) performs the best.

5.2.5 Model development

Proposed defect interaction rule

As aforementioned, the existing rules are not consistent and are deterministic. Therefore, a probabilistic model that holistically considers all the influencing factors including defects configuration and pipe material and geometric properties is developed.

Since the interaction identification response is categorical (i.e., interaction or no interaction), a classification algorithm is suitable to develop the model. In particular, a logistic classification algorithm is adopted, where the binary response is denoted as Y (setting $Y = 0$ for indicating interaction and $Y = 1$ for indicating no interaction), and independent variables are denoted as $X_1 = \{x_i\}$. Thus, the probability of no interaction is expressed as follow:

$$P(Y = 1) = \frac{1}{1 + \exp[-(\beta_0 + \sum_{i=1}^m \beta_i x_i)]} \quad (5.16)$$

where β_i = coefficients for the logistic classification and x_i = pipe properties and adjacent defects characteristics variables. Five normalized variables and their 2nd order interaction among these five variables are used here to construct X_1 , as shown below:

$$X_1 = \left\{ \frac{D}{t}, \frac{\sigma_u}{\sigma_y}, \frac{d_1}{d_2}, N_{S_C} = \frac{S_C}{w_1 + w_2}, N_{S_L} = \frac{S_L}{l_1 + l_2}, 2\text{nd order interaction variables} \right\} \quad (5.17)$$

where d_1, l_1, w_1 = depth, length, and width of Defect 1, respectively; d_2, l_2, w_2 = depth, length, and width of Defect 2, respectively; and N_{S_C} and N_{S_L} = normalized spacing in circumferential and longitudinal direction, respectively. The basic five variables used in X_1 are constructed to reflect all the influencing factors such as defect geometry, colony configurations, and pipe geometrical and material properties. Considering all the 2nd order interaction among the five basic variables, a

total of 20 variables are resulted in X_1 .

When considering all the 20 variables in Eq. (5.16), the model is a full model. A model selection procedure, all possible subset model selection, is adopted to eliminate the ones that do not contribute statistically significantly to the prediction. The models that have any model parameters with p -values greater than 5% are excluded in the model selection. In addition, a maximum model size of five (i.e., five variables in a model) is considered to avoid complex model formulations, and the model performance for each model size is compared using two statistical measures: adjusted R -squared (R^2_{adj}) and mean absolute error (MAE) that is defined as follow:

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (5.18)$$

where y_i = true value, $\hat{y}_i$ = prediction, and n is the number of data points. The model with the highest R^2_{adj} and lowest MAE is the most desirable model for the subset with the same model size (i.e., the same number of variables in the model). Note that when comparing all the best models from all the subsets, these two statistical measures may suggest different models to be the best one. Randomly selected 80% of the data with two defects (called training data) is used for the model selection, while the rest 20% of the data (called test data) is used for the model validation.

After conducting the model selection, it was found that the most desirable model with model size 5 is also the best model overall compared with other model sizes. Table 16 shows the five variable terms selected in this best model and the statistics of the corresponding model parameters in the final model. Note that the proposed interaction rule captures the influencing factors such as the defect depth, length and width, and pipeline strength, thickness and diameter. Also, N_{SC} or N_{SL} is involved in every variable term, indicating normalizing spacing plays a critical role in the interaction prediction, as expected. Using the test data, the proposed model gives a MAE value of 0.3303 that is close to the MAE value of 0.2892 when the training data is used, indicating that the proposed model is valid.

Table 16. Variables and model parameter statistics for the final interaction prediction model

Model Parameters	β_0 (Intercept)	β_1 ($N_{SC} \cdot N_{SL}$)	β_2 ($N_{SC} \cdot D/t$)	β_3 ($N_{SC} \cdot \sigma_u/\sigma_y$)	β_4 ($N_{SC} \cdot d_1/d_2$)	β_5 (N_{SL}^2)	Model Error
Mean	-2.0064	-3.7185	-0.1509	7.6713	-1.7564	3.0116	0
Standard deviation	0.1730	0.5868	0.0255	0.9760	0.5131	0.3082	0.3733

To further evaluate the performance of the proposed model, Figure 89 compares the proposed model (based on MAE and P_{CD} using all the collected data with two defects) with three existing interaction rules, POF [27], API RP 579 [31], and 6WT [30] interaction rules, which found to have better performance than other existing rules shown in Figure 87. Figure 89 shows that the proposed model has the highest P_{CD} and the lowest MAE , indicating that the proposed model has the best performance. This best performance of the proposed model also shows that it is important to holistically include all the influencing factors in the model development.

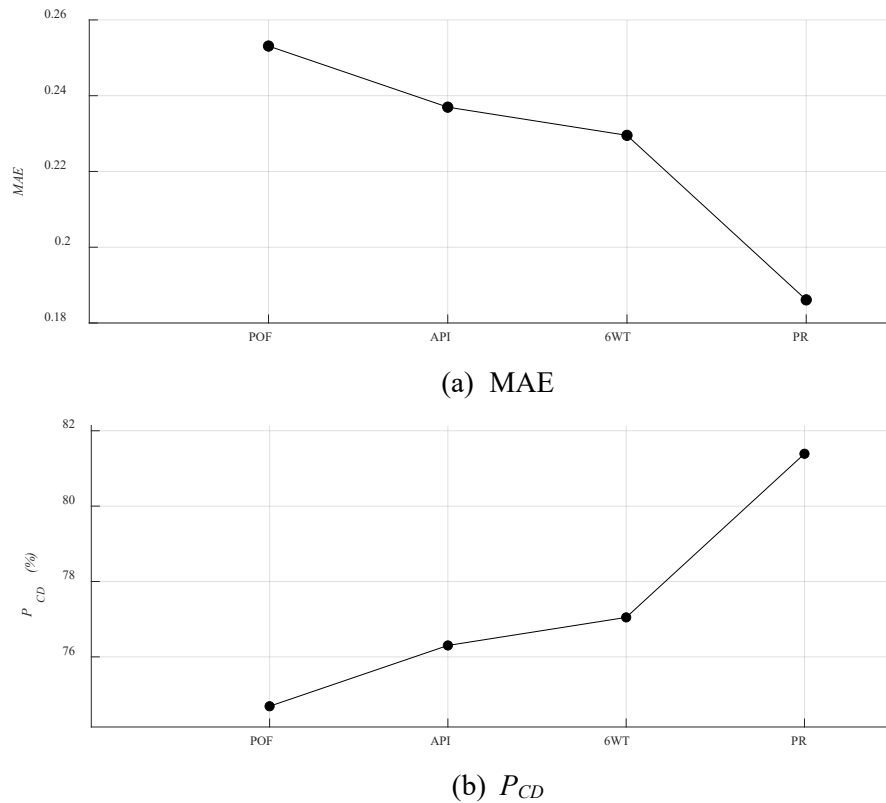


Figure 89. Comparison of the proposed rule (PR) with three best existing interaction rules

Proposed failure pressure prediction model

As shown in Section 4.3.2, the MTI method performs the best compared to the rest of the existing models for failure pressure prediction of pipelines with colony of corrosion defects. However, it is found that in many cases when there is a defect interaction, the pressure predicted by the MTI method is actually the pressure based on a single defect not interacted defects. This means the failure pressure for interacted defects calculated in the MTI method was not able to capture the impact of defect interaction well for those cases. Therefore, a new failure pressure prediction model for interacted defects is proposed by adding a correction factor to the MTI prediction, and

the proposed model is expressed as follow:

$$P_b = \alpha P_{b,MTI} \quad (5.19)$$

$$\alpha = \theta_0 + \sum_{j=1}^m \theta_j x_j + \sigma \varepsilon \quad (5.19a)$$

where $P_{b,MTI}$ = minimum of the failure pressures of all combinations of adjacent defects by the MTI method; and α = correction factor, which is modeled using a multivariate linear regression formulation in which θ_j = model parameters; $X_2 = \{x_j\}$ = independent variables; and $\sigma \varepsilon$ = residual model error in which σ is the standard deviation of the model error (assumed to be constant) and ε is the standard normal random variable (i.e., normality assumption). Five normalized variables and their 2nd order interaction among these five variables are used to construct X_2 , as shown below:

$$X_2 = \left(\frac{D}{t}, \frac{\sigma_u}{\sigma_y}, \frac{d_{clus}}{t}, N_{w_{clus}} = \frac{w_{clus}}{\sqrt{Dt}}, N_{l_{clus}} = \frac{l_{clus}}{\sqrt{Dt}}, \text{2nd order interaction variables} \right) \quad (5.20)$$

where d_{clus} = maximum depth of the combination of adjacent defects with the low failure pressure; l_{clus} and w_{clus} = overall length and width of the combination of adjacent defects with the low failure pressure, respectively. Note that both normalized quantities $N_{l_{clus}}$ and $N_{w_{clus}}$ reflect the defects spacing in longitudinal and circumferential direction, respectively. Considering all the 2nd order interaction among the five basic variables, a total of 20 variables are resulted in X_2 .

The model selection used in Section 5.1 is also applied here to eliminate the variables that do not contribute statistically significantly to the prediction. The maximum model size is chosen to be five and the model performance for each model size is compared using the standard deviation of the model error, σ , which measures the prediction accuracy in the model prediction. The model with the lowest σ is the most desirable model.

The model development is conducted using the data with colony of defects that are identified to have interaction. Table 17 shows the three variable terms selected in the best model and the model parameter statistics in the correction factor, α . As shown in Table 17, all three variable terms include either quantities $N_{l_{clus}}$ or $N_{w_{clus}}$, indicating the colony spacings play an important role in the pressure prediction to account for the defect interaction, as expected.

The performance of the proposed model (PM) for the failure pressure is compared with the

MTI method and Level-1 assessment methods such as the model developed by Kere and Huang (KH) [42] and the DNV RP-F101 method for single defects (Part B) (L1-3) [8]. Using the data with interacted defects, Figure 90 shows the results of the model performance comparison where the cross refers to mean of P_b/P_a and the horizontal lines refer to mean ± 1 standard deviation. Figure 90 indicates that the models L1-3 and KH underestimate the burst pressure, and the MTI method overestimates the burst pressure, while the proposed model provides unbiased prediction. Also, the proposed model has slightly smaller variability in P_b/P_a compared to the MTI method. Therefore, one can conclude that the correction factor proposed improves the MTI method for calculating the failure pressure prediction of a pipe with interacted defects.

Table 17. Variables and model parameter statistics for the correction factor

Model parameters	θ_0 (Intercept)	θ_1 $(N_{w_{clus}} \cdot d_{clus}/t)$	θ_2 $(N_{l_{clus}} \cdot d_{clus}/t)$	θ_3 $(N_{w_{clus}}^2)$	Model Error, $\sigma\epsilon$
Mean	1.1087	-0.1885	-0.0369	0.0112	0
Standard deviation	0.0085	0.0091	0.0038	0.0009	0.0918

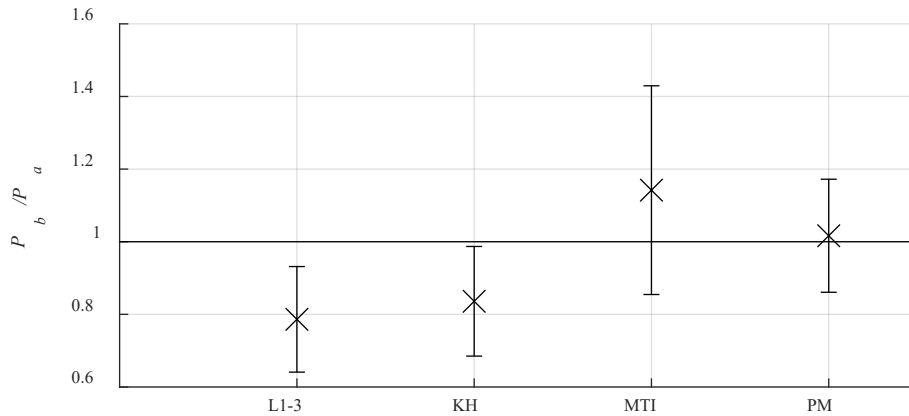


Figure 90. Comparison of the proposed model (PM) with the MTI method and two Level-1 assessment methods

5.3 Models for Pipeline with Single Crack-like Defect

5.3.1 Background

For pipelines with crack-like defects, the burst failure pressure has been assessed using the “pipeline specific” methods and “generic” methods [43]. The “pipeline specific” methods are developed for the pipeline industry and include models like the original Ln-Sec [44], modified Ln-Sec [45], and CorLASTM [46]. The “generic” methods are standards such as API 579 [28] and BS 7910 [29] using failure assessment diagram (FAD). In fact, different formulas are used to assess the failure pressure of pipelines with crack-like defects depending on the thickness of the pipeline (i.e., thin-walled or thick-walled pipeline), the extend of the defect (i.e., through-wall or part-wall defect), and the location (i.e., internal or external surface) and orientation (i.e., inclined, longitudinal or circumferential) of the crack. For example, the Battle Memorial Institute developed the equations called “NG-18 equations” for pipelines subjected to only internal pressure with longitudinally oriented through-wall and part-wall defects [44]. Staat [47] proposed prediction equations for thick-walled pipes containing axial cracks. Also, model like PRCI MAT-8 [48] were developed to assess the failure pressure for pipelines with longitudinal seam weld cracks. The API 579 [28] and BS 7910 [29] provide formulations for different levels of assessment accuracy for thin-walled and thick-walled pipelines that contains through-wall and part-wall crack-like defects oriented longitudinally or circumferentially on the internal or external surface of the pipeline.

To evaluate the performance of the existing models, several studies have been conducted. For instance, Tandon et al. [49] compared three models (i.e., modified Ln-Sec, CorLASTM, and API 579-version 2007) using the ratios of the actual to the predicted burst failure pressures based on 15 full scale burst test data. The results showed that the modified Ln-Sec and CorLAS models both have an average error of about 7% and the API 579 about 22%. Also, Hosseini [50] compared the performance of four models (i.e., original Ln-Sec, CorLAS, BS 7910-version 2005, API 579-version 2007) using the percent error of predictions based on 4 full scale burst test data. The CorLAS model was found to be the best model and the BS 7910 model was the most conservative. In addition, Yan et al. [51] performed model error comparison of five models (i.e., original Ln-Sec, CorLAS, BS 7910-version 2005, API 579-version 2007 and R6-Rev4 Amendment 10 [52]) using the ratios of the actual to the predicted burst failure pressures based on 112 full scale burst test data. Their results showed that the CorLAS model has the best performance, and the original Ln-Sec, BS 7910, and API 579 are in general conservative, which is not suitable to be used in the

risk assessment of pipelines.

In this Section 5, firstly, a comprehensive database (consists of experimental and numerical data, a total of 160 data sets) is established, which consists of the data collected from the literature and the numerical data obtained from validated finite element models conducted in this study using the extended finite element method (XFEM). Then, a performance comparison of five existing prediction models (i.e., original Ln-Sec, modified Ln-Sec, CorLAS, API 579 and BS 7910) is conducted. Next, the proposed failure pressure model is developed by adding a correction factor to an existing model (i.e., the modified Ln-Sec model that is identified to be the best existing model in terms of prediction accuracy); and the correction factor is constructed using a multivariate linear regression fitted by the database established. Lastly, a life cycle cost analysis of a pipeline with a single crack like defect is conducted to evaluate the impact of failure prediction models on the expected total life cycle cost of pipelines.

5.3.2 Existing prediction models

In the literature, the most models used to predict the failure pressure model of pipelines with single crack-like defects are the original Ln-Sec [44], modified Ln-Sec model [45], CorLAS [46], and failure assessment diagram (FAD) methods such as API 579 [28] and BS 7910 [29]. These models are described in this section.

Ln-Sec Model

Ln-Sec (log-secant) model also known as NG-18 equation is a semi-empirical model developed by the Battle Memorial Institute in the late 1960s to predict the burst pressure of pipes containing longitudinally oriented surface cracks subjected to only internal pressure [44]. The burst pressure, P_b is computed using the minimum of the values using two criteria, the flow stress and fracture toughness dependent criteria, and can be expressed as,

$$P_b = \min \left[\frac{2t\sigma_f}{D} \cdot \frac{1 - a/d_w}{1 - (a/d_w)M_T^{-1}}, \frac{4d_w\sigma_f}{\pi D} \cdot \frac{1 - a/d_w}{1 - (a/d_w)M_T^{-1}} \arccos \left(e^{-\frac{\pi K_{mat}^2}{8c_{eq}\sigma_f^2}} \right) \right] \quad (5.21)$$

$$M_T = \begin{cases} \sqrt{1 + 0.6275 \frac{(2c_{eq})^2}{Dd_w} - 0.003375 \frac{(2c_{eq})^4}{(Dd_w)^2}}, & \frac{(2c_{eq})^2}{Dd_w} \leq 50 \\ 3.3 + 0.032 \frac{(2c_{eq})^2}{Dd_w}, & \frac{(2c_{eq})^2}{Dd_w} > 50 \end{cases} \quad (5.21a)$$

$$\sigma_f = \sigma_y + 68.95 \text{ MPa} \quad (5.21b)$$

$$2c_{eq} = A/a \quad (5.21c)$$

where D = outside diameter of the pipe, d_w = wall thickness of the pipe, $2c_{eq}$ = equivalent length of crack, a : depth of crack, A = actual area of the surface crack along its length, σ_f = flow stress of the pipe material, σ_y = yield strength of the pipe material, M_T = folias or bulging factor of pipe, and K_{mat} = fracture toughness of the pipe material. If actual value of K_{mat} is not available, it can be approximated using the following empirical expression:

$$K_{mat}^2 = \frac{C_v E}{A_c} \quad (5.21d)$$

where C_v = upper shelf energy determined from tests of Charpy V-notch impact specimens, A_c = cross-sectional area of the Charpy specimen used, and E = Young's modulus of the pipe material.

Modified Ln-Sec Model

As the original Ln-Sec model underestimates the failure pressure for long and shallows defects, it was then modified by Kiefner [45] in 2008 by adding a correction factor. The modified Ln-Sec model formulation is given by:

$$P_b = \frac{2t\sigma_f}{D} \cdot \frac{1 - A/A_0}{1 - (A/A_0)M_T^{-1}} \cdot \frac{\arccos\left(e^{-\frac{\pi K_{mat}^2}{8c_{eq}\sigma_f^2}}\right)}{\arccos\left(e^{-\frac{\pi K_{mat}^2}{8c_{eq}\sigma_f^2} \cdot \frac{1}{1 - \left(\frac{a}{l}\right)^{0.8}}}}\right)} \quad (5.22)$$

where A_0 = reference area = $2c \cdot d_w$ in which $2c = l$ = crack length

CorLASTTM

CorLASTTM is a software-based model developed to assess axially oriented crack-like surface flaws subjected to only internal pressure using elastic-plastic fracture mechanics [46]. This model applies two failure criteria to evaluate the burst pressure: one criterion is based on flow strength and the other one is based on the fracture toughness of the pipe material. The burst pressure is assessed by:

$$P_b = \min(\sigma_f, \sigma_l) \cdot \frac{2t}{D} \cdot \frac{1 - A/A_0}{1 - (A/A_0)M_T^{-1}} \quad (5.23)$$

Where M_T can be calculated using Eqs. (5.21a), σ_l = local stress at the tip of the crack when failure occurs. The flow stress, σ_f , can be evaluated by either Eq. (5.21b) or the following equation that is

generally recommended:

$$\sigma_f = \frac{\sigma_y + \sigma_u}{2} \quad (5.23a)$$

where σ_u = ultimate strength of the pipe material. The local stress, σ_l , is calculated iteratively by setting applied integral J equal to the fracture toughness of the pipe material, J_c , and J depends on σ_l and the corresponding function is given as:

$$J = Q_f F_{sf} a \left[\frac{\sigma_l^2 \pi}{E} + f_3(n) \varepsilon_p \sigma_l \right] \quad (5.23b)$$

where Q_f = elliptical flaw shape factor, F_{sf} = free-surface factor, $f_3(n)$ = function of strain-hardening exponent, n = strain-hardening exponent, and ε_p = plastic strain. The formulation to calculate J_c can be found in [46].

FAD models

FAD models are developed based on fracture mechanics and can be used to assess failure pressure of pipeline with crack-like defects with two key parameters: the load ratio, L_r , and the toughness ratio, K_r , which are calculated using the following expressions [28]:

$$L_r = \frac{\sigma_{ref}}{\sigma_y} \quad (5.24)$$

$$K_r = \frac{K_I}{K_{mat}} \quad (5.25)$$

where σ_{ref} = reference stress and K_I = stress intensity factor [7].

Figure 91 is modified based on [28] and illustrates the concept of the FAD models. The failure pressure is determined by finding the pressure that causes the assessment point (L_r , K_r) to fall on the assessment or cut-off line [51]. In this study, the models based on this FAD principle are the API RP 579 (version 2016) [28] and the BS 7910 (version 2013) [29], which define the assessment line differently.

API RP 579

The API RP 579 model includes three assessment levels to evaluate the failure pressure depending on the available information on material and geometry properties of the pipeline and operating conditions [28]. Specifically, Level 1 assessment uses limited information on pipe properties, crack geometry, and operations conditions such as the length and depth of crack, operating temperature, and loads. Level 2 assessment requires further detailed information on material properties, loading

conditions, and the state of stress at the location of the flaw. Level 3 assessment requires the maximum information on material properties and crack geometry such as stress-strain curve, geometry and material dependent FAD, and crack growth; thus, Level 3 assessment usually can provide most accurate failure predictions. However, considering the information available for the data collected in this study, Level 2 assessment is considered for the model prediction evaluation. The assessment line function for Level 2 is defined as:

$$K_r = (1 - 0.14L_r^2)[0.3 + 0.7\exp(-0.65L_r^6)] \quad \text{for } L_r < L_{r(\max)} \quad (5.26)$$

$$L_{r(\max)} = \frac{\sigma_y + \sigma_u}{2\sigma_y} \quad (5.26a)$$

where $L_{r(\max)}$ defines the cut-off line. Then the assessment point is calculated using Eqs (5.24) and (5.25) with values of σ_{ref} , K_I , and K_{mat} which can be estimated through [28].

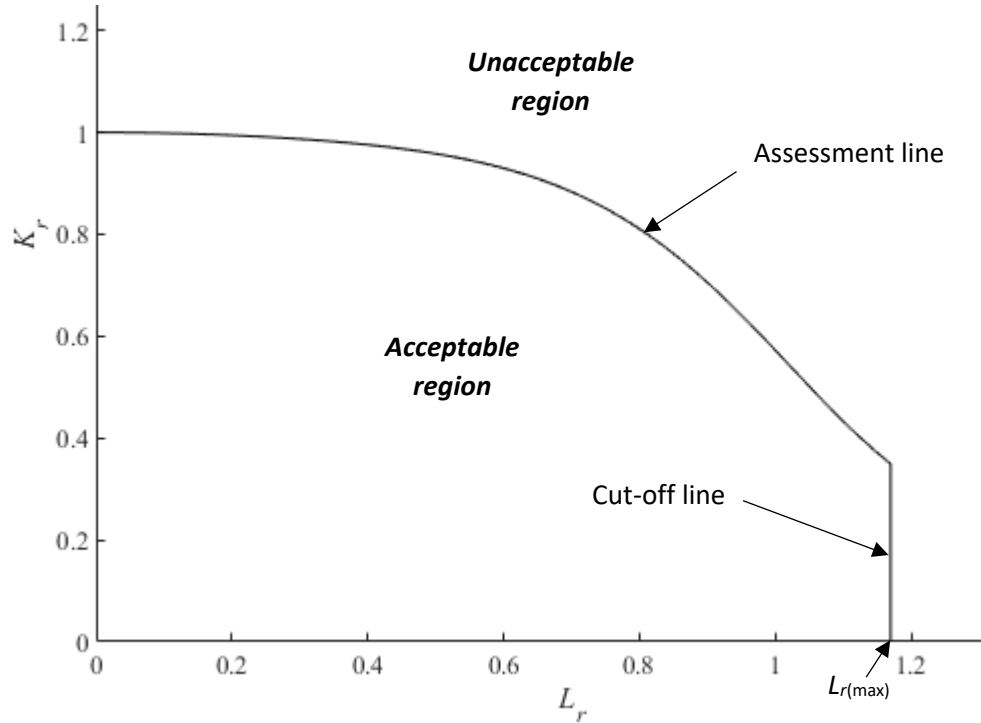


Figure 91. Illustration of the FAD

BS 7910

The BS 7910 have three options to evaluate the failure pressure based on the application and materials data available [29]. Option 1 is a conservative procedure that does not require the stress-strain curve. Option 2 uses a material-specific stress-strain curve. Option 3 generates a FAD using

numerical analysis and is not limited to materials showing ductile tearing. In this study, BS 7910 Option 1 is used since the stress-strain curve is not always available. The followings equations are used to describe the assessment line function:

$$K_r = \begin{cases} \left(1 + \frac{1}{2} L_r^2\right)^{-1/2} \cdot [0.3 + 0.7 \exp(-\mu L_r^6)], & L_r \leq 1 \\ K_{r,1} \cdot L_r^{(N-1)/(2N)}, & 1 < L_r < L_{r(\max)} \\ 0, & L_r \geq L_{r(\max)} \end{cases} \quad (5.27)$$

$$\mu = \min\left(0.001 \frac{E}{\sigma_y}, 0.6\right) \quad (5.27a)$$

$$N = 0.3 \left(1 - \frac{\sigma_y}{\sigma_u}\right) \quad (5.27b)$$

$$L_{r(\max)} = \frac{\sigma_y + \sigma_u}{2\sigma_y} \quad (5.27c)$$

where $K_{r,1}$ refers to K_r by setting $L_r = 1$. Then the assessment point is calculated using Eqs. (5.24) and (5.25) with values of σ_{ref} , K_I , and K_{mat} which can be estimated through [29].

5.3.3 Data Collection

A comprehensive failure pressure data with longitudinally oriented single crack-like defect is established in this study, which is used to evaluate the performance of the existing prediction models, and also to develop failure pressure assessment method. The database established consists of the data directly collected from literature and additional numerical data obtained from FE analysis conducted in this study.

Data collected from literature

A total of 122 different laboratory experimental burst test results of thin-walled pipes (i.e., $D/d_w \geq 20$) with external longitudinal oriented single crack are collected from literature. Table 18 provides a summary of the data collected. Since the failure pressure of a pipe with crack-like defect depends on the pipe material and geometry properties, and defect geometry, it is worthy to examine the correlation of the following quantities with the failure pressure (P_b) through scatter plots shown in Figure 92 using the data collected: yield strength (σ_y), Charpy shelf energy (C_v), ratio of pipe diameter to pipe thickness (D/d_w), ratio of defect depth to pipe thickness (a/d_w), and ratio of defect half-length to pipe thickness (c/d_w).

Figure 92(a), the scatter plot of σ_y vs. P_b , shows that an overall increase of σ_y leads to a higher burst pressure as expected. Figure 92(b) on the other hand indicates that a low D/d_w ratio

(referring to very thick-wall pipes) leads to a high burst pressure, except for a few cases circled by the dotted lines. These cases may be explained by their values of crack depth close to half of pipe thickness as shown in the dotted circle in Figure 92(c). Figure 92(c) indicates a negative correlation between P_b and a/d_w , while Figure 92(d) shows a positive correlation between P_b and a/c . Figure 92(e) shows the scatter plot of C_v over P_b , and no distinct trend is shown.

Table 18. Summary of the database collected from literature

Grade	D/d_w	σ_y (MPa)	σ_u (MPa)	C_v (J)	a/d_w	a/c	Counts
X52	[48.70 – 94.9]	[341 – 456]	[487 – 627]	[21.69 – 42.03]	[0.24 – 0.92]	[0.02 – 0.34]	19
X60	[32.15 – 103.96]	[379 – 510]	[536 – 634]	[27.12 – 135.00]	[0.19 – 0.77]	[0.02 – 0.24]	22
X65	[66.80 – 100.48]	[363 – 514]	[525 – 656]	[10.85 – 132.87]	[0.25 – 0.87]	[0.0032 – 0.14]	12
X100	[55.76 – 73.89]	[739 – 795]	[171 – 261]	[10.85 – 132.87]	[0.19 – 0.55]	[0.02 – 0.12]	4
15Mo3	22.23	246	570	84	[0.78 – 0.95]	[0.07 – 0.38]	11
34CrMo4	[26.08 – 32.29]	[703 – 878]	[874 – 990]	[59 – 81]	[0.51 – 0.99]	[0.08 – 0.48]	16
4134V	[30.46 – 34.05]	[1048 – 1096]	[1138 – 1179]	[21.69 – 27.12]	[0.62 – 0.96]	[0.13 – 0.51]	19
St 35	22.23	336	486	76	[0.25 – 0.90]	[0.02 – 0.24]	13
St 70	[86.26 – 86.73]	[529 – 543]	[670 – 695]	[50 – 115]	[0.68 – 0.95]	[0.05 – 0.09]	6

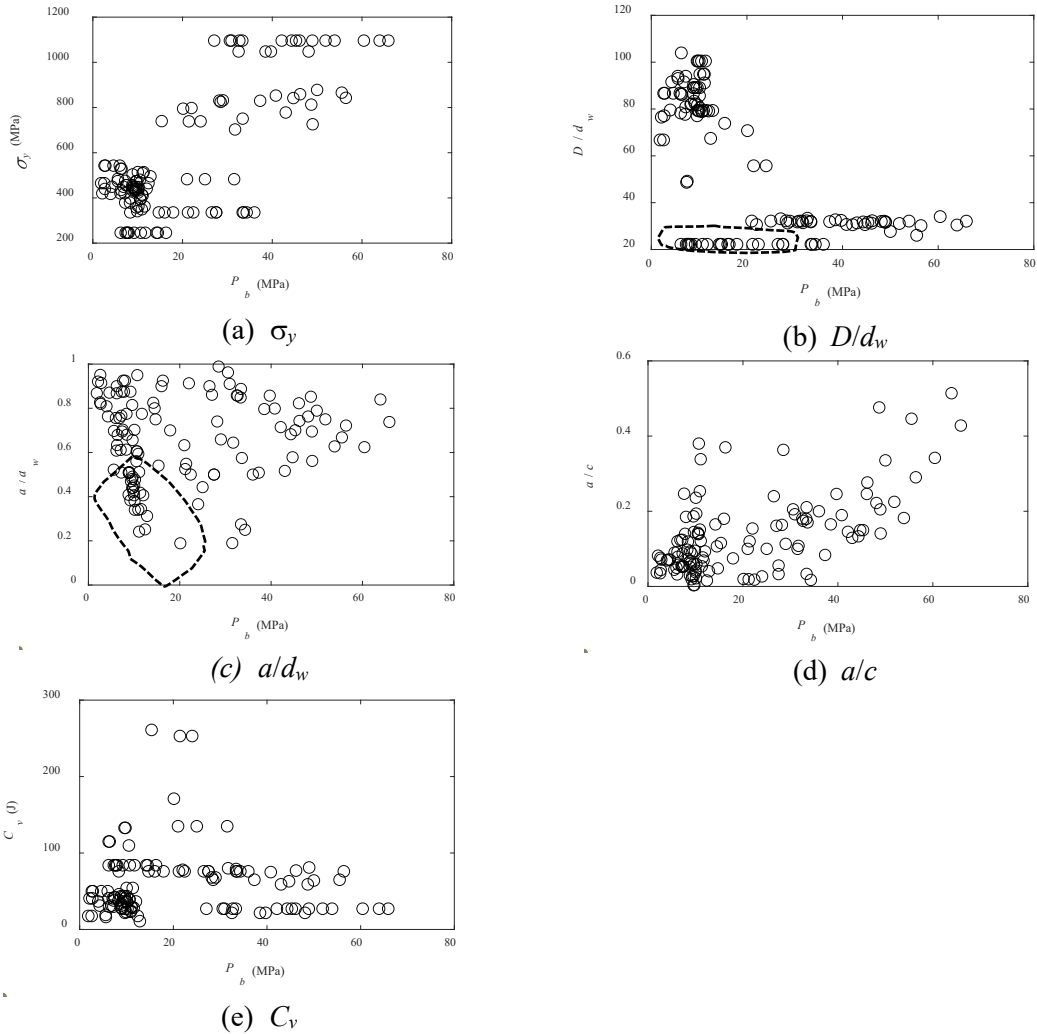


Figure 92. Scatter plots of burst pressure (P_b) vs. selected quantities

Additional numerical data

To complement the existing data collected in the literature, FEMs are used to generate additional data points. In this study, the FEMs are developed in ABAQUS. While modeling cracking growth using the conventional FE methods is challenging due to the need of mesh conformity to the geometry discontinuities and remeshing as crack grows, the extended finite element method (XFEM) has been developed to addresses these challenges [53]. The XFEM was introduced by Belytschko and Black [54] and it extends the conventional FEM through using the partition of unity property of finite elements by adding enriching degrees of freedom with special displacement functions to the finite element approximation.

In Abaqus/Standard, XFEM uses a traction-separation model for crack propagation

consisting of a crack initiation criterion and a damage propagation law. The crack initiation criteria available in Abaqus/Standard [53] are the maximum principal stress (*Maxps*), the maximum principal strain (*Maxpe*), the maximum nominal stress (*Maxs*), the maximum nominal stress (*Maxe*), the quadratic nominal stress (*Quads*), and the quadratic nominal strain (*Quade*). The available damage propagation criteria are fracture energy (G_c) and displacement of crack tip at failure. In the literature, the *Maxps* or *Maxpe* and G_c criteria are the most used criteria to model crack propagation in XFEM. For example, Lin et al. [55] used the *Maxps* and G_c criteria to simulate crack propagation in pressurized steel pipes; and Okodi et al. [56] used the *Maxpe* and G_c criteria to predict the burst pressure of longitudinally cracked pipelines. In fact, the *Maxpe* criterion can be represented as follows [53]:

$$f = \left\{ \frac{\langle \varepsilon_{max} \rangle}{\varepsilon_{max}^o} \right\} \quad (5.28)$$

where f = maximum strain ratio, ε_{max} = maximum principal strain, ε_{max}^o = maximum allowable principal strain, and the Macaulay brackets indicates that a purely compressive strain does not cause damage initiation. When f reaches a value of one, damage is assumed to initiate. The fracture energy G_c is the energy required to create a unit of crack area. Since *Maxpe* and G_c are material properties, which can be obtained by calibration using the burst test results obtained from literature. With the obtained *Maxpe* and G_c , this material is then used to generate additional XFEM models of pipelines with longitudinally oriented single crack-like defect to determine the failure pressure.

For computational efficiency, only half of the pipe (i.e., mid length of pipe) with appropriate boundary conditions is modeled due to the symmetry condition, as shown in Figure 93. Also, the crack is modeled to have semi elliptical shape. The model is meshed using 8-node fully integrated linear brick elements (C3D8) for the solid part and linear quadrilateral elements (S4R) for the shell part; and elasto-plastic materials are used. Furthermore, Statics-General procedure is used for the analysis and an internal pressure is gradually applied until the pipe fails. Failure is determined when the crack propagation reaches the last elements of the wall thickness of the pipe. Figure 94 illustrates the crack propagation in one of the selected cases: the crack propagates until the last element of the wall thickness of the pipe, which causes the pipe to fail.

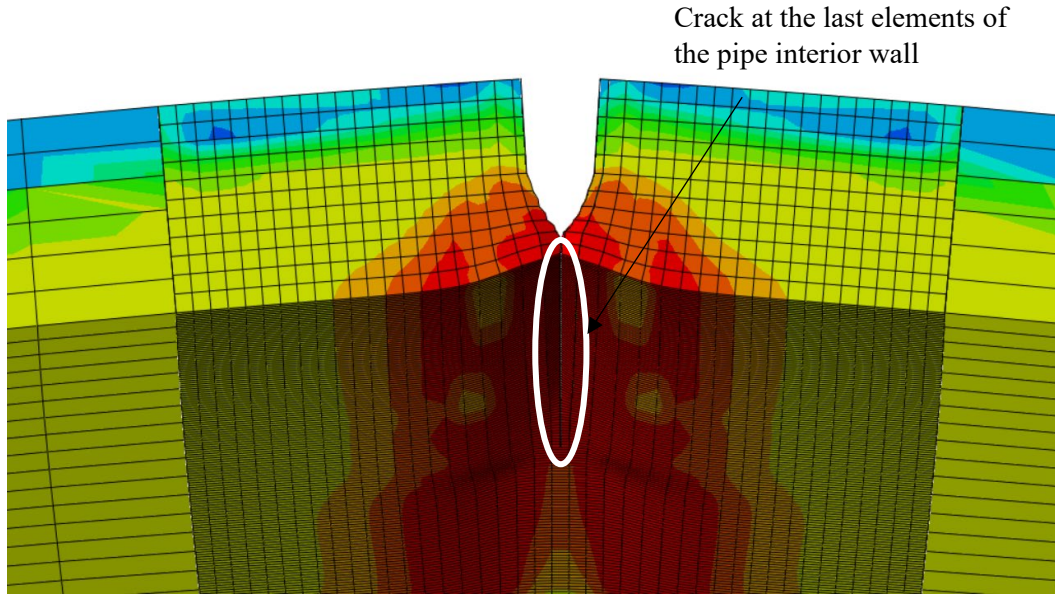


Figure 93. Crack propagation in a pipe modeled in Abaqus

To calibrate the parameters $Maxpe$ and G_c , the J integral of the material, J_{mat} , is calculated first using the J integral for linear elastic material under mode loading I expressed as follows [57]:

$$J_{mat} = \frac{K_{mat}^2}{E'} \quad (5.29)$$

where $E' = E$ for plane stress condition and $E' = E/(1 - \nu^2)$ for plane strain condition in which ν = Poisson's ratio of the material. Using the calculated J_{mat} (a measure of the fracture toughness of the material) as a starting value for G_c (referring to fracture energy), $Maxpe$ and G_c are calibrated by trial and error using a few experimental testing data collected from the literature. Table 19 summarizes the pipe material properties and defect geometries of the experimental testing [58]. As shown in Table 19, there are four different materials, except X52 the other three materials have four different cases. For each material, $Maxpe$ and G_c are calibrated so that the difference between the failure pressure obtained from the FE analysis ($P_{b,FE}$) conducted in this study and the failure pressure data reported in the literature ($P_{b,test}$) is small as much as possible. Table 19 shows that with the calibrated $Maxpe$ and G_c , the error percentages, $(P_{b,test} - P_{b,FE})/P_b$, are all within 10% for all cases. Thus, $Maxpe$ and G_c are obtained for the selected materials, and the FE models with these four materials are validated. It is worth stating that the accuracy of the XFEM parameters depends on the number of burst test data used for calibration. Okodi et al. [59] assert that there is inherent error in calibrating damage parameters using few burst test results. However, because of the

limitation of burst test data available in the literature, the damage parameters can only be obtained based on the limited data available.

Table 19. Experimental testing cases selected for calibrating $Maxpe$ and G_c and calibrated results

Grade	D (mm)	t (mm)	σ_y (MPa)	σ_u (MPa)	J_{mat} (N/mm)	a (mm)	$2c$ (mm)	$P_{b,test}$ (MPa)	$P_{b,FE}$ (MPa)	Error (%)	$Maxpe$	G_c (N/mm)
X52 [49]	508	6.4	350	497	50.63	3.8	30	10.92	10.98	-0.55	0.02	50
X60 [50]	508	5.70	433	618	43.5	2.17	200.00	10.10	10.18	-0.79	0.084	50
						2.68	200.00	9.30	9.06	2.58		
						2.74	200.00	9.60	8.90	7.29		
						2.91	200.00	8.83	8.65	2.04		
4134V [58]	237	7.0	1096	1179	62.16	4.34	201.00	60.42	66.01	-9.25	0.07	50
	236	7.4				5.44	25.40	65.89	68.43	-3.85		
	237	7.6				5.72	50.80	51.88	52.64	-1.46		
	237	7.4				6.27	69.85	33.32	34.40	-3.24		
4134V [58]	236	7.4	1048	1138	38.70	5.64	50.80	48.10	48.29	-0.40	0.07	40
	236	7.3				6.25	50.80	39.70	41.72	-5.09		
	237	7.2				5.77	51.80	38.50	35.95	6.62		
	237	7.3				6.30	69.85	32.50	32.53	-0.09		

With the damage parameters of these four materials, 38 new numerical cases are generated. These added numerical cases are designed to cover a wider range of three quantities: D/t , a/d_w , and a/c . Figure 94 shows the scatter plots of σ_y vs the three quantities based on the experimental and FE burst test data collected from literature and the new generated numerical cases. As shown in Figure 94, the new cases are designed to cover the regions where the data collected from literature is scarce. The pipeline properties and defect geometries and FE results ($P_{b,FE}$) of these 38 new cases are provided in Table 20.

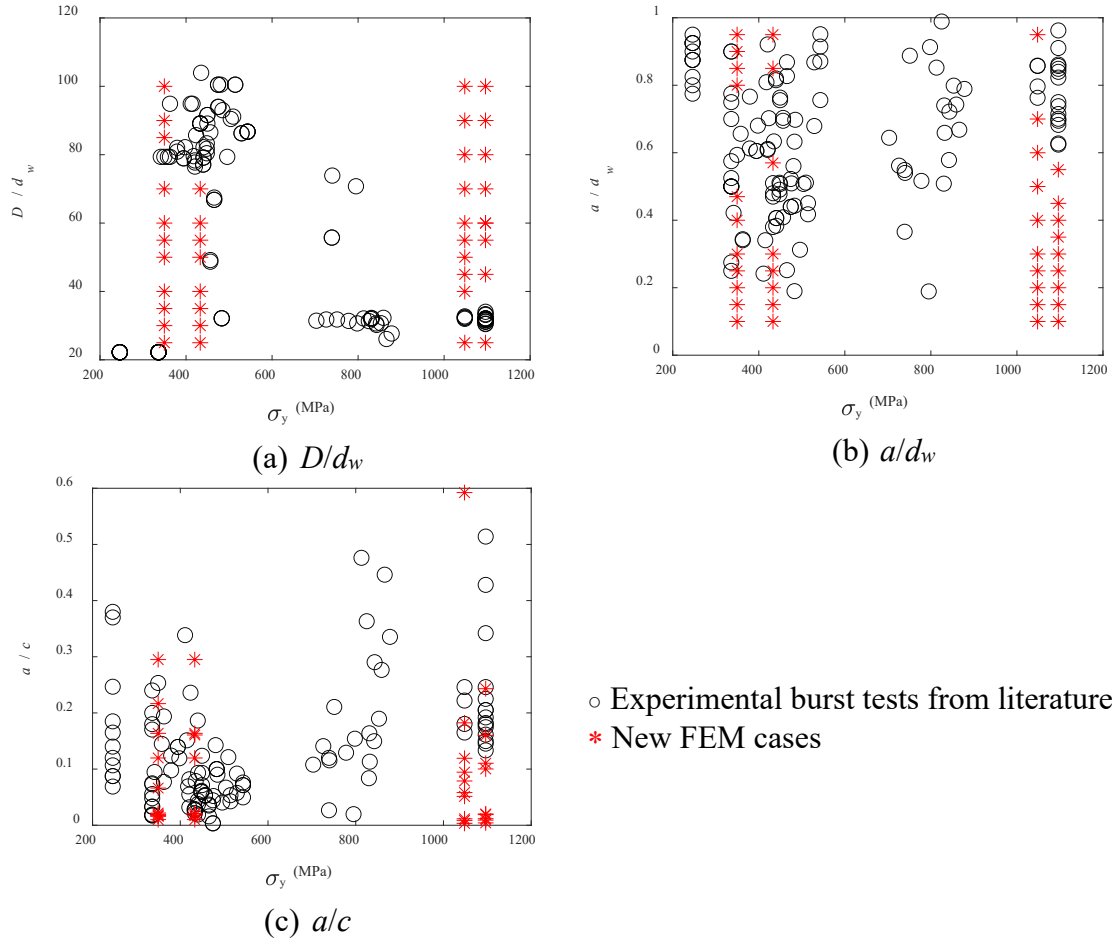


Figure 94. Scatter plots of selected quantities vs. yield stress (σ_y)

Table 20. New FE models cases

Grade	D (mm)	t (mm)	σ_y (MPa)	σ_u (MPa)	CVN (J)	d (mm)	$2c$ (mm)	$P_{b,FE}$ (MPa)
X52	250.00	8.33	350	497	23.05	0.83	10.18	38.41
X52	400.00	8.00	350	497	23.05	1.60	26.72	21.54
X52	315.00	7.88	350	497	23.05	2.36	472.54	20.90
X52	450.00	7.50	350	497	23.05	3.53	429.27	10.78
X52	508.00	7.26	350	497	23.05	1.09	100.11	15.07
X52	300.00	8.57	350	497	23.05	2.14	14.53	30.97
X52	215.00	8.60	350	497	23.05	7.31	863.68	6.85
X52	425.00	7.73	350	497	23.05	7.34	896.43	1.83
X52	762	7.62	350	497	23.05	6.86	207.13	2.09
X52	600	6.67	350	497	23.05	5.33	49.26	6.99
X52	510	6.00	350	497	23.05	2.40	247.91	8.93
X60	250.00	8.33	433	618	43.50	0.83	10.18	44.67
X60	400.00	8.00	433	618	43.50	1.60	26.72	25.24
X60	315.00	7.88	433	618	43.50	2.36	472.54	24.50
X60	450.00	7.50	433	618	43.50	4.28	429.27	10.42
X60	508.00	7.26	433	618	43.50	1.09	13.55	19.02
X60	300.00	8.57	433	618	43.50	2.14	14.53	36.34
X60	215.00	8.60	433	618	43.50	7.31	863.68	9.22
X60	425.00	7.73	433	618	43.50	7.34	698.14	2.41
4134V	508	5.64	1048	1138	21.69	0.56	11.95	29.01
4134V	615	8.79	1048	1138	21.69	1.76	44.58	34.33
4134V	315	3.94	1048	1138	21.69	1.18	12.96	30.56
4134V	250	5.56	1048	1138	21.69	2.22	37.27	49.28
4134V	350	8.75	1048	1138	21.69	4.38	150.43	38.55
4134V	300	6.00	1048	1138	21.69	3.60	12.16	47.96
4134V	215	8.60	1048	1138	21.69	8.17	317.73	11.22
4134V	400	6.67	1048	1138	21.69	1.00	231.43	40.66
4134V	275	5.00	1048	1138	21.69	1.25	744.79	38.27
4134V	425	4.25	1048	1138	21.69	2.98	517.76	9.15
4134V	762	9.53	1096	1179	27.12	0.95	19.01	33.64
4134V	350	7.78	1096	1179	27.12	2.33	19.19	58.05
4134V	215	8.60	1096	1179	27.12	3.44	43.00	96.21
4134V	335	5.58	1096	1179	27.12	1.12	117.56	39.23
4134V	450	5.00	1096	1179	27.12	0.75	13.59	29.38
4134V	475	8.64	1096	1179	27.12	2.16	473.26	37.96
4134V	515	5.15	1096	1179	27.12	2.32	230.81	16.45
4134V	220	3.14	1096	1179	27.12	1.10	528.15	24.97
4134V	236	3.93	1096	1179	27.12	2.16	371.17	22.02

Performance comparison of existing models

The performance comparison of the existing failure prediction models considered in this study is conducted using the mean, standard deviation of the ratio of the predicted to the actual burst failure pressures, P_b/P_a . Note that only the original Ln-Sec and modified Ln-Sec models are applicable to all the data points (i.e., 160 data points). The CorLAS model is only applicable to data that meets the application restriction listed in Table 21, which corresponds to 136 data points. For the API RP 579 and BS 7910 models, the equations for calculating σ_{ref} and K_I in this study are only applicable to data satisfying the application restriction listed in Table 21, resulting that only 93 data points are applicable for API 579 and 143 data points for BS 7910. Therefore, to be fair for all the models, the comparison is performed using the common data points (i.e., 81 data points) that are applicable to all the models. Also, it is worth stating that instead of using the software CorLAS™, the method used by the software was directly used to calculate the failure pressure. Figure 95 shows the performance comparison of the five existing models, where the crosses refer to mean of P_b/P_a and the horizontal bars refer to mean ± 1 standard deviation. As shown in Figure 95, all prediction models (except Corlas) averagely underestimate the failure pressure (shown as cross markers below the horizontal line of 1.0). Figure 95 also indicates that most prediction models have big variability, particularly API RP579 model. One could also observe that the modified Ln-Sec model has the best performance with the smallest bias and variation. Furthermore, Figure 95 shows that the API RP579 and BS 7910 models are overly conservative, and this can be understood by the fact that the models based on FAD are developed to avoid failure rather than failure prediction [51].

Table 21. Conditions for using the CorLAS , API 579 and BS7910 models in this study

Models	Application constrain	Crack and geometry dimensional limits
CorLAS	Pipeline with $\sigma_y/E < 0.005$	-
API 579	Cylinder subjected to internal pressure and containing longitudinally oriented surface cracks with semi-elliptical shape	$0 \leq a/d_w \leq 0.8$ $0.03125 \leq a/c \leq 2.0$ $0 \leq d_w/(D/2 - d_w) \leq 1.0$
BS 7910		$0 \leq a/d_w \leq 0.8$ $0.05 \leq a/c \leq 1.0$ $0.1 \leq d_w/(D/2 - d_w) \leq 0.25$

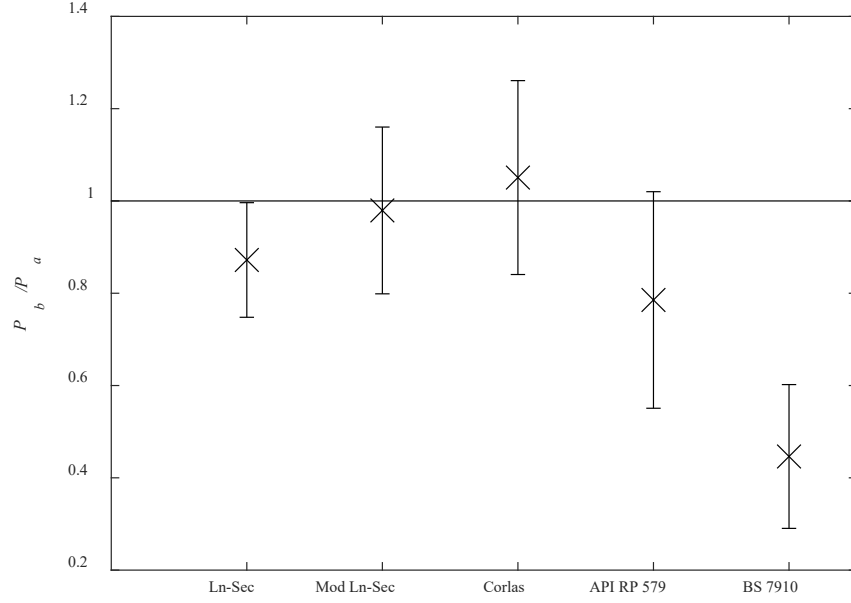


Figure 95. Comparison of existing failure pressure models

5.3.4 Proposed model development

As shown in the previous section, the modified Ln-Sec model performs the best compared to the rest of the existing models for failure pressure prediction of pipelines with single like-crack defects. Here, the proposed failure pressure, P_b , is modeled by adding a correction factor, α , to the modified Ln-Sec model, $P_{Mod Ln-Sec}$, to improve the model prediction accuracy. The proposed model can be expressed as follows:

$$P_b = \alpha \cdot P_{Mod Ln-Sec} \quad (5.30)$$

This correction factor, α , is modeled using a multivariate linear regression formulation in this study as follows:

$$\alpha = \beta_0 + \sum_{i=1}^m \beta_i x_i + \sigma \varepsilon \quad (5.31)$$

where β_i = model parameters; $\mathbf{X} = \{x_i\}$ = independent variables; and $\sigma \varepsilon$ = residual model error in which σ is the standard deviation of the model error (assumed to be constant) and ε is the standard normal random variable (i.e., normality assumption). Four normalized variables and their 2nd order interaction among these four variables are used here to construct $\mathbf{X}$, as shown below:

$$\mathbf{X} = \left(\frac{D}{d_w}, \frac{\sigma_u}{\sigma_y}, \frac{a}{d_w}, \frac{a}{c}, \frac{c}{d_w}, \frac{l}{\sqrt{Dd_w}}, 2^{\text{nd}} \text{ order interaction} \right) \quad (5.32)$$

Considering all the 2nd order interaction among the four basic variables, a total of 27 variables are resulted in X . When considering all the 27 variables in Eq. (5.31), the model is a full model. An all-possible-subset model selection is adopted to eliminate the ones that do not contribute statistically significantly to the prediction [25]. In addition, a maximum model size of five (i.e., five variables in a model) is considered to avoid complex model formulations, and the model performance for each model size is compared using the model error standard deviation, σ . The model with the lowest σ is the most desirable model.

The model development is conducted using the database established in this study. After model selection, it is found that the model with size 5 is the best model overall compared with other sizes models. Table 22 shows the variables selected and the statistics of the corresponding model parameters in the final model. It is worth noting that the data ranges used for the model development are D/d_w in [22 100], a/d_w in [0.10 0.99], and a/c in [0.0032 0.5140]. The prediction performance of the proposed model is then compared with the modified Ln-Sec model through the mean, standard deviation of the ratio of the predicted to the actual burst failure pressures, P_b/P_a , using all the data (i.e., 160 data points) as shown in Figure 96, where the cross refers to mean and the horizontal lines refer to mean $\pm$ 1 standard deviation. Figure 96 indicates that the modified Ln-Sec model and the proposed model provides unbiased prediction. However, the proposed model shows smaller variability in P_b/P_a . Therefore, one can conclude that the correction factor proposed improves the modified Ln-Sec model accuracy.

Table 22. Variables and model parameter statistics for the correction factor

Model Parameters	β_0 Intercept	β_1 a/d_w	β_2 $D/d_w \cdot a/d_w$	β_3 $D/d_w \cdot c/d_w$	β_4 $\sigma_u/\sigma_y \cdot a/d_w$	β_5 $(\sigma_u/\sigma_y)^2$	Model Error
Mean	0.889	1.205	-1.841*10 ⁻³	5.665*10 ⁻⁵	-1.280	0.256	0
Standard deviation	0.111	0.334	6.377*10 ⁻⁴	6.443*10 ⁻⁶	0.263	0.0656	0.124

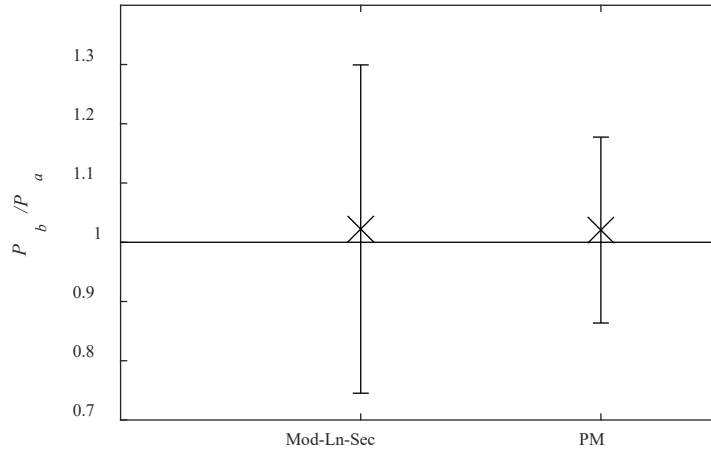


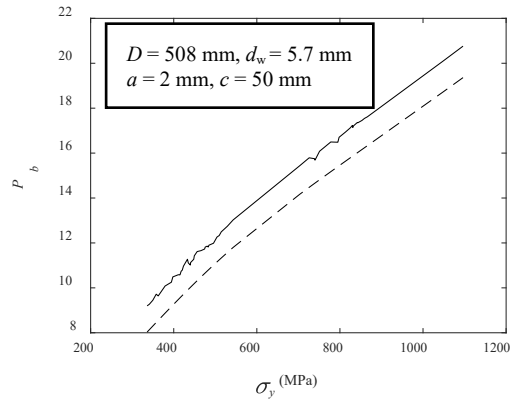
Figure 96. Comparison of P_b/P_a using the modified Ln-Sec model (Mod Ln-sec) and the proposed model (PM)

5.3.5 Sensitivity analysis

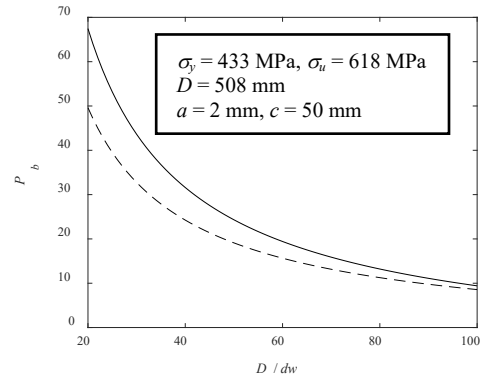
To obtain a better understanding of how material and geometry properties and crack sizes impact the failure pressure, a sensitivity analysis is performed to evaluate the impact of four selected quantities (i.e., σ_y , D/d_w , a/d_w , and c/d_w) on both the proposed model and the modified Ln-Sec model, for a comparison purpose. Figure 97 shows the changes in the pressure prediction of a pipeline (that has $\sigma_y = 433$ MPa, $\sigma_u = 618$ MPa, $D = 508$ mm, $d_w = 5.7$ mm, $a = 2$ mm, and $c = 50$ mm) when varying D/d_w , a/d_w , and c/d_w .

Overall, as shown in the four plots Figure 97, the predicted burst pressure are sensitive to all four quantities, showing the importance of these quantities in the model. Secondly, the trends of the change in P_b over the quantities are similar for both models, meaning that adding the correction factor in the proposed model does not fundamentally change the relationship between those quantities to the prediction.

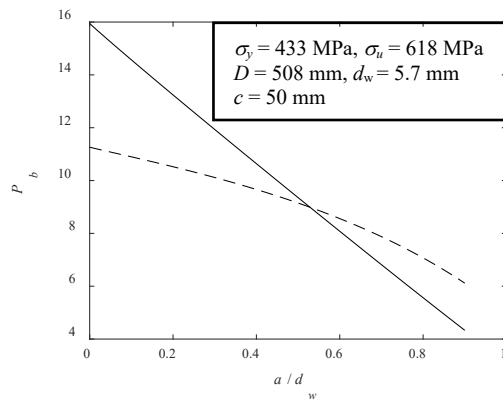
As expected, the failure pressure predictions of both models increase with the increase of σ_y as shown in Figure 97 (a); and the change rates are the same for both models. Meanwhile, Figure 97(b), Figure 97(c) and Figure 97(d) show that the predictions of both models decrease with the increase of D/d_w , a/d_w , and c/d_w ; however, the change rate in the proposed model is higher (particularly for a/d_w), meaning the proposed model is more sensitive to the wall thickness ratio and crack depth. In addition, the predicted pressure from the proposed model is higher than the existing model expect when $a/d_w > 0.5$ as shown in Figure 97(c).



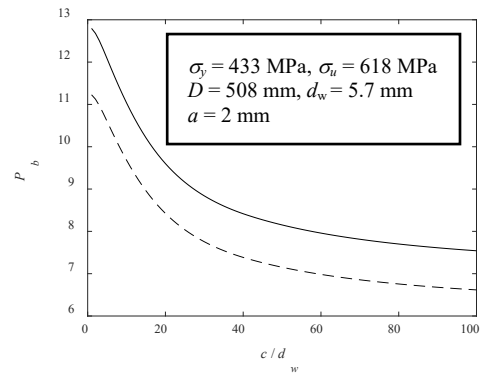
(a)



(b)



(c)



(d)

— Proposed model prediction - - - - - Modified Ln-Sec prediction

Figure 97. Sensitivity of the four selected quantities on pressure prediction

6 Objective 4: Uncertainty Impact on Pipeline Reliability

6.1 Background

A pressurized pipeline generally fails in two distinctive modes: small leak (when a corrosion defect penetrates the pipe wall thickness) or a burst (when the operating pressure of the pipe exceeds the burst pressure of the pipe) [35]. To evaluate the impact of the proposed prediction models on the pipeline structural performance, the probability of the burst failure of a pipeline with corrosion defect using the proposed models is evaluated. The probability of failure, P_f , is defined as the conditional probability of attaining or exceeding prescribed limit states given a set of boundary variables, and can be written as:

$$P_f = \int_{g(\mathbf{X}) \leq 0} f(\mathbf{X}) d\mathbf{X} \quad (6.1)$$

where $f(\mathbf{X})$ is the joint probability density function of a vector of random variables, $\mathbf{X}$; $g(\mathbf{X})$ is limit-state function; and $g(\mathbf{X}) \leq 0$ refers to the failure domain. This probability is assessed by conducting a reliability analysis such as Monte Carlo simulations and First/Second Order Reliability Methods (FORM/SORM). The limit-state function for a burst failure is defined as follow:

$$g(\mathbf{X}) = P_b - P_p \quad (6.2)$$

where P_b is the pressure capacity of the pipe (that is usually estimated by the failure pressure prediction model), and P_p is the demand (that is the operating pressure of the pipe). In practice, reliability index is calculated to measure the pipe performance, and a generalized reliability index is defined as [36]:

$$\beta = \Phi^{-1}(1 - P_f) \quad (6.3)$$

where Φ^{-1} refers to the inverse of cumulative distribution function of standard normal distribution.

6.2 Case Study I

In this session, the reliability index is calculated based on the failure pressure prediction based on either the proposed models (shown in Table 8) or the best existing model (i.e., G5-19) for pipelines with single corrosion defect. The random variables, $\mathbf{X}$, used in the reliability analysis and their distribution information are listed in Table 23. The model errors of the best existing models are calculated using the established database so that they could be accounted in the reliability analysis. Comparing the model errors in Table 23 indicates that the best existing models are biased

(especially for Level 3) and have larger standard deviations of model errors compared to the proposed models, consistent with the observations in Figure 80.

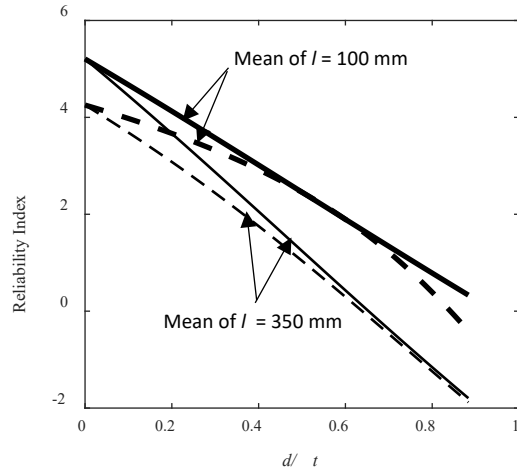
Table 23. Distribution parameters of random variables used in Case Study I

Random variable	Distribution	COV (%)	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
Outside diameter of pipe, D (mm)	Normal	5	324			16.2		
Nominal wall thickness, t (mm)	Normal	5	6			0.3		
Defect depth, d (mm)	Normal	5	-			-		
Defect length, l (mm)	Normal	5	100 or 350			5 or 17.5		
Yield strength, σ_y (MPa)	Normal	3	357	534	589	10.71	16.02	17.67
Ultimate strength, σ_u (MPa)	Normal	3	458	661	731	13.74	19.83	21.93
Operating Pressure, P_p (MPa)	Normal	5	7.61	11.39	12.57	0.38	0.57	0.63
Model error in the proposed model (MPa)	Normal	-	0	0	0	1.84	1.07	1.36
Model error in the best existing model (MPa)	Normal	-	0.39	-0.53	0.90	2.23	1.39	2.45

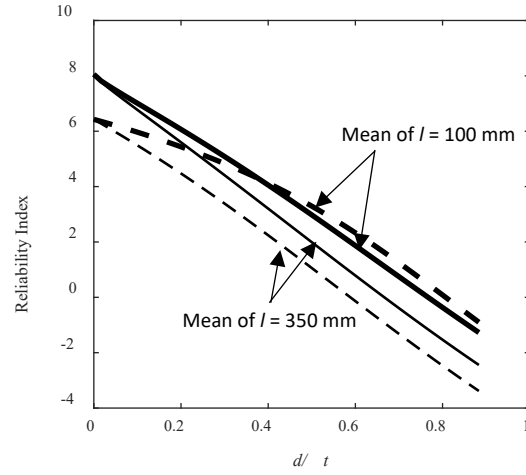
Figure 98 compares the reliability index of pipelines with the failure pressure capacities calculated based on either the proposed models or the best existing models for three materials (corresponding to three levels of σ_u) under various levels of corrosion depths (d/t varying from 0% to 90%) and two levels of defect lengths (means of $l = 100$ mm and 350 mm). As expected, the reliability index decreases with the increase of the defect depth on the pipe for a given defect length; the reliability index is smaller for longer defect length. Also, the defect length impact becomes more significant when the corrosion depth increases. This indicates that both corrosion depth and length are critical particularly when corrosion becomes worse.

When comparing the reliability curves resulted from the two pressure prediction models (solid lines vs. dashed lines in Figure 98), one can notice that the reliability based on the proposed models is higher than the one based on the best existing model, especially for smaller d/t . Such difference is bigger for the defect with $l = 350$ mm and also bigger for material with Level 3 σ_u . For example, Figure 98(c) shows that for a defect depth of 40% of wall thickness and a defect length of 350 mm the reliability index based on the proposed model and best existing model are about 3.9 ($P_f = 0.000048$) and 2.3 ($P_f = 0.0107$), respectively, which is substantially different. In this case, using the best existing model may lead to unnecessary costs associated with repairs and maintenance prompted by the lower reliability index calculated. This result indicates that the failure pressure prediction model plays an important role in the reliability evaluation of a pipeline;

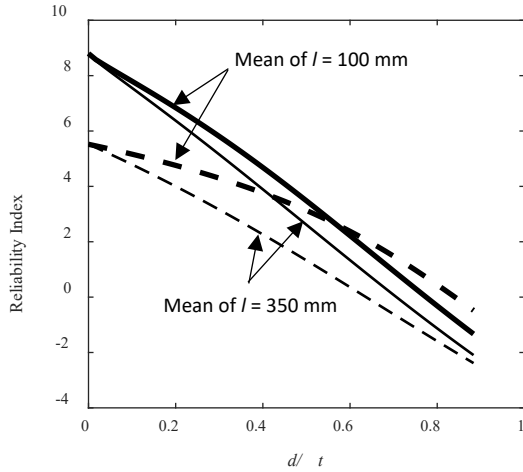
the difference resulted from the proposed models and the best existing models is not negligible.



(a) Material with Level 1 σ_u



(b) Material with Level 2 σ_u

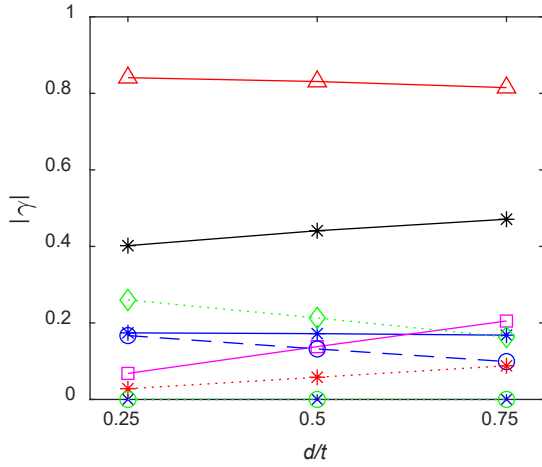


(c) Material with Level 3 σ_u

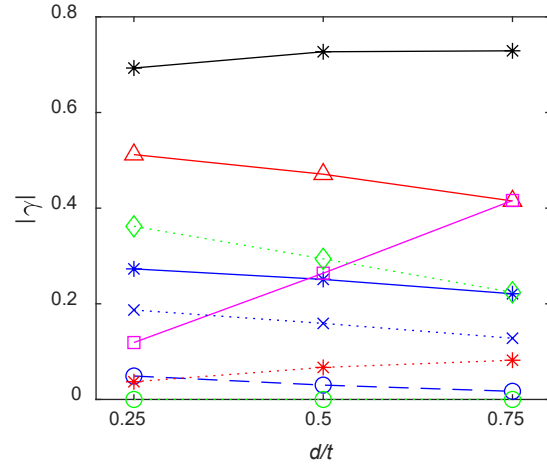
— Proposed model
- - - Best existing model

Figure 98. Defect depth-dependent reliability index based on the proposed models and best existing models for pipelines with single corrosion defect

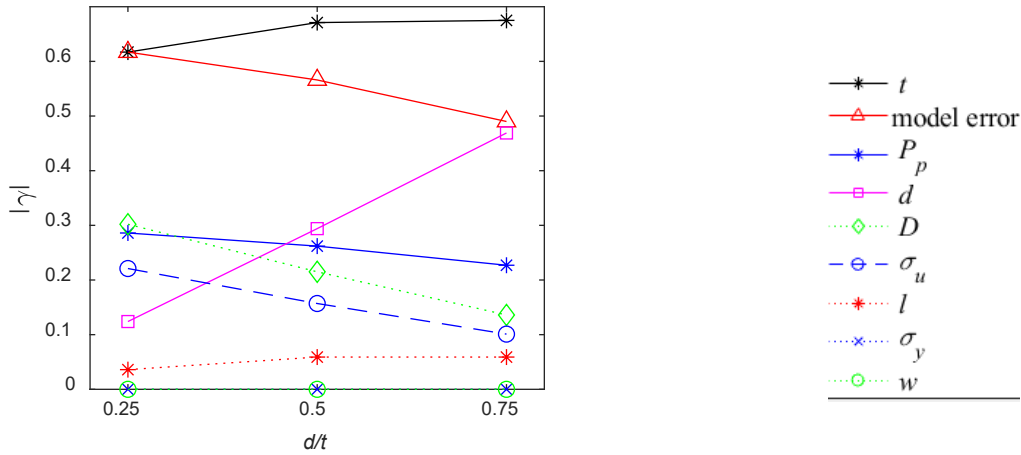
In addition, importance analysis [62] are used to identify important uncertainty sources that contribute most to the variability of the pipeline performance. Figure 99 shows the absolute value of importance measure, γ , of each random variable that considered in the reliability analysis. For all three scenarios, the top two most important variables are wall thickness (t) and model error in the burst pressure model. This indicates the majority of uncertainties stem from wall thickness and burst pressure prediction. In addition, the plots in Figure 99 also indicate the defect depth becomes more important when corrosion becomes more severe, as expected. Lastly, the defect length and width are among the least important variables.



(a) Material with Level 1 σ_u



(b) Material with Level 2 σ_u

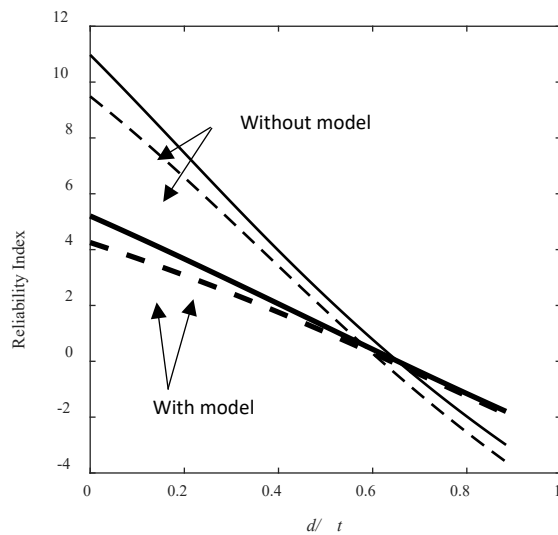


(c) Material with Level 3 σ_u

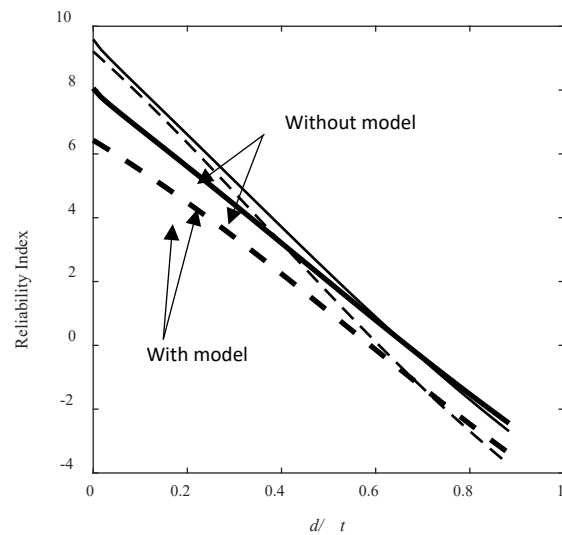
Figure 99. Absolute values of importance measures based on the proposed models for pipelines with single corrosion defect assuming $l = 100\text{mm}$

To evaluate the impact of the model error (that reflects the model accuracy) in the performance evaluation, Figure 100 compares the reliability curves with and without considering the model errors for the same three pipe materials in Figure 98 under various levels of corrosion depths (d/t varying from 0% to 90%) and one level of defect length (mean of $l = 350\text{ mm}$). Similar to Figure 98, the reliability index difference between the solid line (obtained using the proposed model) and the dashed line (obtained using the best existing model) is substantial particular for material with Level 3 σ_u , as shown in Figure 100(c). For either the proposed or best existing models, it is apparent that the model error has a great impact on the reliability index; and considering model error has a much lower index value. For instance, the reliability index calculated

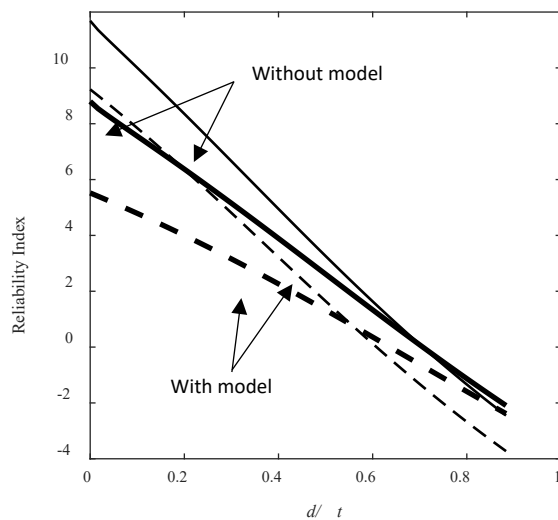
based on the proposed model for the pipeline with Level 3 σ_u and a defect depth of 20% are about 8.4 and 6.4 without and with model error respectively. This shows that if the model error is not considered, the overestimated reliability index can cause a delay of the pipeline maintenance and repair, leading to unexpected pipeline failure with tremendous consequences (both economically and environmentally). Therefore, it is important to include the model error in the reliability evaluation of corroded pipelines. In summary, based on Figure 98 and Figure 100, one can conclude that the performance of a failure pressure prediction model plays a critical role in determining the structural performance of corroded pipelines.



(d) Material with Level 1 σ_u



(e) Material with Level 2 σ_u



(f) Material with Level 3 σ_u

— Proposed model
 - - - Best existing model

Figure 100. Defect depth-dependent reliability index calculated with and without the model error for the proposed models and best existing models for pipelines with single corrosion defect

6.3 Case Study II

To evaluate the impact of defects interaction on the structural integrity of pipelines, the time-dependent probability of the burst failure of a pipeline with two corrosion defects is evaluated. The pressure capacity of the pipeline with the colony of defects is calculated based on an interaction rule and failure pressure prediction models. When defect interaction is identified, the proposed model for interacted defects is applied. When there is no defect interaction identified, the pressure capacity is determined by the smallest value of all pressures based on each single defect within the colony using the failure pressure equation in the MTI method, which is also the DNV RP-F101 method for single defects (Part B). Note that the interaction rules (that are based on defect geometries) and the failure pressure prediction models are time-dependent, as they depend on defect sizes. Thus, the probability of failure is time-dependent,

To predict the defect size time-evolution, the corrosion growth model developed by Caley et al. [63] is considered here, and it is written as:

$$d_m(t) = k(t - t_0)^\gamma \quad (6.4)$$

where $d_m(t)$ = average value of the maximum pit depth at time t ; t_0 is the corrosion initiation time; and k and γ are the pitting proportionality and exponent factors, respectively. Considering a mixed type of soil, the value of k and γ are estimated to be 0.164 mm/year and 0.780, respectively [63]. On the other hand, no defect length or width growth models are available; thus, the corrosion defect length and width growths are simply calculated using a corrosion defect length to depth ratio and a corrosion defect width to depth ratio, respectively.

The random variables, $\mathbf{X}$, used in the reliability analysis and their distribution information are listed in Table 24. Figure 101 compares the time-dependent reliability index of a pipeline with the failure pressure capacities calculated based on various interaction rules: the proposed interaction rule (PR) developed in this study and three existing interaction rules (i.e., KV [26], DNV RP F101 [8], and 6WT [30]). For a comparison purpose, Figure 101 also shows the reliability index curve of the pipeline when only one defect is considered: dashed line for considering Defect 1 only and black solid line for considering Defect 2 only.

As expected, the reliability index decreases with time since defects grow with time. The reliability index curve based on the interaction rules KV or 6WT overlaps with the one that

considers only Defect 2, which indicates that the interaction rules KV and 6WT do not recognize any defect interaction during the 75-year time window. Meanwhile, the interaction rule DNV identifies the defect interaction around year 40, and then the corresponding reliability index dropped from a value of 6.0 ($P_f = 9.87 \times 10^{-10}$) to a much lower level, 3.7 ($P_f = 1.08 \times 10^{-4}$), due to the identified interaction. On the other hand, the proposed rule (PR) identifies the defect interaction much earlier, around year 4, where the reliability index curve drop occurs accordingly. Figure 101 clearly shows that the interaction effect plays a critical role in the time-dependent performance evaluation of a pipeline with colony of corrosion defects. In this case study, using the existing interaction rule KV, 6WT or DNV does not recognize the interaction at all or at a much later time, which may cause a delay of the pipeline maintenance and repair leading to unexpected pipeline failure usually with both adverse consequences economically and environmentally.

Table 24. Distribution parameters of random variables used in Case Study II

Random variable	Distribution	COV (%)	Mean	Standard deviation
Outside diameter of pipe, D (mm)	Normal	5	324	16.2
Nominal wall thickness, t (mm)	Normal	5	6	0.3
Yield strength, σ_y (MPa)	Normal	3	534	16.02
Ultimate strength, σ_u (MPa)	Normal	3	661	19.83
Operating pressure, OP (MPa)	Normal	5	14	0.7
Corrosion defect length to depth ratio	-	-	5	-
Corrosion defect width to depth ratio	-	-	2	-
Defect 1 corrosion initiation time, $t_{0,d1}$ (year)	-	-	3	-
Defect 2 corrosion initiation time, $t_{0,d2}$ (year)	-	-	2	-
Initial longitudinal spacing, $S_{L,init}$, between defects (mm)	-	-	100	-
Initial circumferential spacing, $S_{C,init}$, between defects (mm)	-	-	50	-

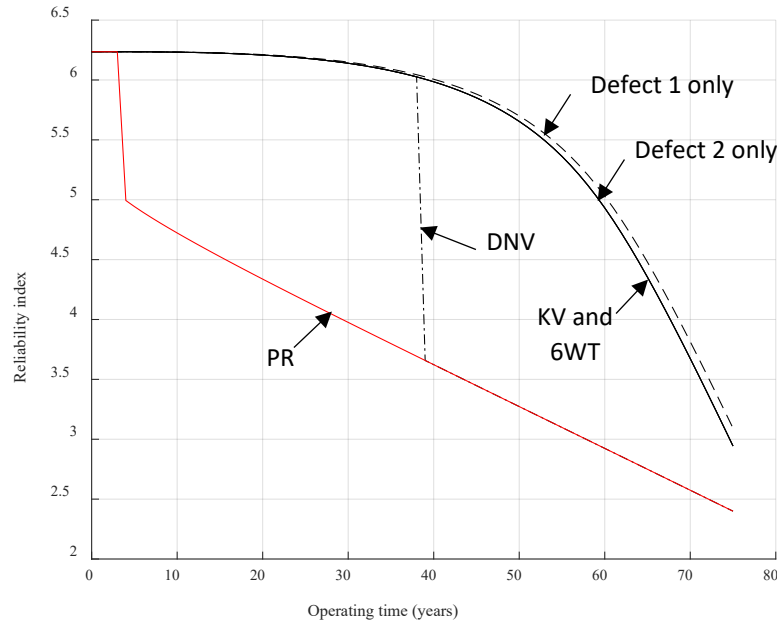


Figure 101. Reliability index versus operating time based on different interaction rules

6.4 Case Study III

In this section, a case study of a pipeline with a crack-like defect is considered to evaluate the impact of failure prediction models on the life-cycle cost. The framework of expected life-cycle cost developed by Kere and Huang [64] is used in this study. The framework is developed based on a decision tree model with the use of analytical methods to evaluate events. The expected total life cycle cost consists of cost of inspection, repair, and failure with the consideration of discount rate and service life. Using the total probability concept, the expected total life cycle cost is determined by adding each conditioned expected total cost on a given number of failures occurrence during the service life multiplied by the probability of the corresponding failure occurrences, and is written as:

$$E[C_T] = \sum_{k=0}^n E[C_{T,k}] = \sum_{k=0}^n P_{f,k} \cdot E[C_T | k \text{ failures}] \quad (6.5)$$

where $C_{T,k}$ = cost due to the scenario when k failures occur during the service life, $P_{f,k}$ = probability of k number of failures occurrence during the service life, and $E[C_T | k \text{ failures}]$ = expected total cost given k number of failures occurrence. In particular, $P_{f,k}$ and $E[C_T | k \text{ failures}]$ are calculated using the formulations given in [64], where the probabilities of different numbers of failure occurrences are probabilistically assessed using probability distribution of failure time with the

consideration of the impact of repair actions that are possibly taken after each planned inspection, and the probability distribution of failure time is determined based on the time-dependent probability of failure. The probability of failure is defined as in Eq. (6.2). Note that the failure pressure prediction model is time-dependent, as it depends on defect size. Thus, the probability of failure is time-dependent.

To predict the defect size time-evolution due to fatigue loading, the Paris law proposed by Paris and Erdogan [65] is considered in this study, which is expressed as follows:

$$\frac{da}{dN} = C(\Delta K)^m \quad (6.6)$$

where a = crack size; N = number of cycles; C and m = material constants; and ΔK = stress intensity range. The stress intensity range is given by [66]

$$\Delta K = Y(a)\Delta\sigma\sqrt{\pi a} \quad (6.6a)$$

where Y = geometry factor depending on the geometry of the crack, a , and $\Delta\sigma$ = stress range. For simplicity, Y is assumed to be constant. Using an iterative process with an initial crack size, a_0 , the defect size time-evolution, $a(t)$, can be predicted. For this case study, the parameters used in the crack growth model are listed in Table 25.

Figure 102(a) shows the crack defect growth over time, t , for this case study, where the solid line refers to the prediction with mean values of coefficient and the dotted line refers to the mean prediction ± 1 standard deviation. Figure 102(a) indicates that the variability in the defect depth prediction is very small for lower value of t , but starts to become significant with increase in t . Next, the burst pressure capacity is evaluated using the proposed failure pressure model (PM) and the modified Ln-Sec model (Mod Ln-Sec) separately, and the predicted failure pressure, $P_b(t)$, is shown in Figure 102(b). With the same material and geometry properties used by Hosseini et al [50] as shown in Table 6 and an operating pressure assumed as a normal random variable with a mean of 8 MPa and standard deviation of 0.4 MPa, the probability of burst failure, $P_f(t)$, can be assessed through Eq. (6.2) using reliability analysis. Figure 102(c) shows the resulted $P_f(t)$ using PM or Mod Ln-Sec for the failure pressure prediction. For a better visualization of the impact of the pressure prediction model on the probability of failure, $P_f(t)$ is also plotted in the logarithmic scale as shown in Figure 102(d). As expected, Figure 102(b) shows that the failure pressure prediction using either of the models decreases with the increase of t , while the probability of failure increases with t as shown in Figure 102(c) and Figure 102(d), since the defect depth

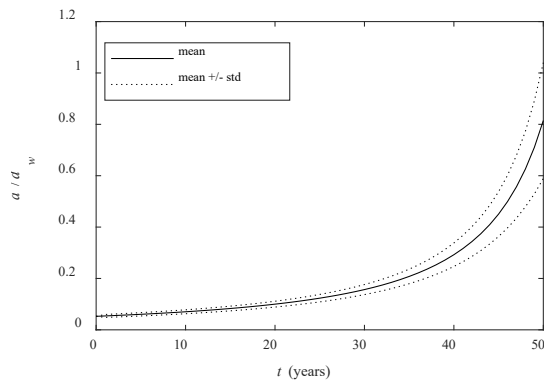
increases with time. Moreover, Figure 102(b) shows that Mod Ln-Sec is not sensitive to the defect growth for $t < 30$ years in this case study; and the predicted pressure from PM is much higher than the one from Mod Ln-Sec especially when $t > 48$ years. Because of the difference in the failure pressure prediction (that is the failure pressure prediction from PM is much higher than the one from Mod Ln-Sec), Figure 102(d) shows that the probability of failure using PM is much smaller compared to the one using Mod Ln-Sec model.

For the life cycle cost analysis, the unit costs for inspection (C_I), repair (C_R), and failure (C_F) are assumed: $C_I = a_{in} \cdot C_0$, $C_R = a_r \cdot C_0$, and $C_F = a_f \cdot C_0$, where $a_{in} (= 0.0177)$, $a_r (= 0.243)$, and $a_f (= 100)$ are multiplicative factors for inspection, repair, and failure, respectively. Those factor values are chosen based on the ranges presented in Gomes and Beck [67]. Also, the discount rate is assumed to be 2%, considering a service life of 50 years. For simplicity, we assume that the inspection interval, Δt , is fixed and set to be 10 years, and the defect repair threshold, d_r , is the variable. Specifically, d_r is set to be within the range of [15–30%] of the wall thickness, d_w . Using Eq. (6.5) by setting $n = 2$ (i.e., ignoring the consequence due to 3 or more failure occurrences), the expected total cost, $E[C_T]$, is calculated. Figure 103 displays $E[C_T]$ and its three components $E[C_{T,0}]$, $E[C_{T,1}]$, and $E[C_{T,2}]$ for different values of d_r/d_w considering PM and Mod Ln-Sec in the probabilities of failures evaluation. It found that the lowest $E[C_T]$ value shown in red circle in Figure 103(a) occurs at $d_r/d_w = 0.23$ and 0.20 using PM and Mod Ln-Sec, respectively. The optimal defect repair threshold, d_r , using Mod Ln-Sec is smaller since the probability of failure using Mod Ln-Sec is bigger (as shown in Figure 102(d)), and lower d_r is preferred for higher probability of failure to increase the chance of performing maintenance actions, which can prevent event failure. For both models PM and Mod Ln-Sec, Figure 103(b), shows that the expected total cost given no failure occurrence, $E[C_{T,0}]$, decreases with the increase of d_r/d_w , since the number of possible repairs decreases with the increase of d_r/d_w . $E[C_{T,0}]$ using PM is higher because the probability of no failure is higher for PM. The expected total cost given one failure, $E[C_{T,1}]$, shown in Figure 103(c) and the expected total cost given two failures, $E[C_{T,2}]$, shown in Figure 103(d) increase with d_r/d_w , because the probability of failure using both models increase due to the lower number of possible repair actions when setting a higher repair criterion. Figure 103(c) indicates that $E[C_{T,1}]$ using PM can be neglected for $d_r/d_w < 0.23$ and Figure 103(d) shows that although $E[C_{T,2}]$ using Mod Ln-Sec is higher compared to the one using PM, $E[C_{T,2}]$ is negligible for both models. The results from Figure 103 clearly indicate that the failure pressure model considered in the life cycle

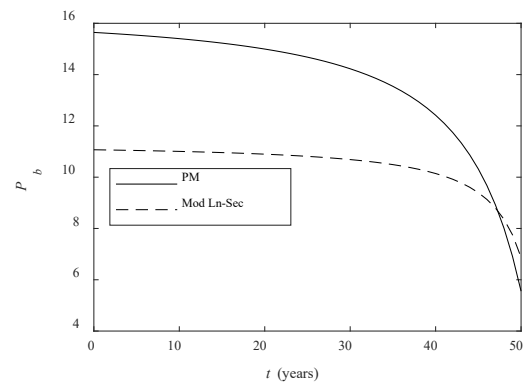
cost analysis influences the decision making in the risk management.

Table 25. Distribution parameters of random variables used in Case Study III

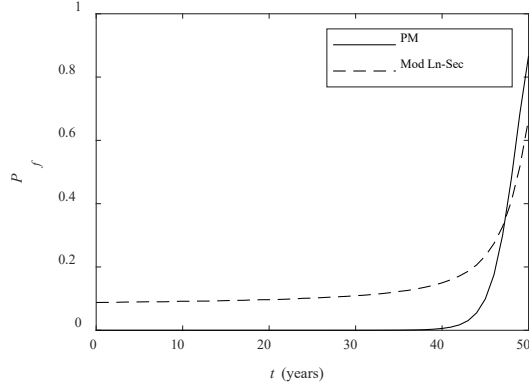
Random variable	Distribution	Mean	STD
Outside diameter of pipe, D (mm)	Normal	508 [50]	25.4
Nominal wall thickness, d_w (mm)	Normal	5.7 [50]	0.285
Yield strength, σ_y (MPa)	Normal	433 [50]	12.99
Ultimate strength, σ_u (MPa)	Normal	618 [50]	18.54
Estimated fracture toughness, K_{mat} (MPa ² ·m)	Normal	335.49	16.77
Material parameter, C	lognormal	$2.3 \cdot 10^{-12}$ [68]	$6.9 \cdot 10^{-13}$ [68]
Material parameter, m	Normal	3.0 [68]	0.3 [68]
Geometry function, Y	-	1	-
Stress range, $\Delta\sigma$ (MPa)	Weibull	14	1.4
Number of load cycles per year, N	-	10^6	-
Initial crack depth, a_0 (mm)	Normal	0.3	0.03
Crack length, $2c$ (mm)	-	100	-



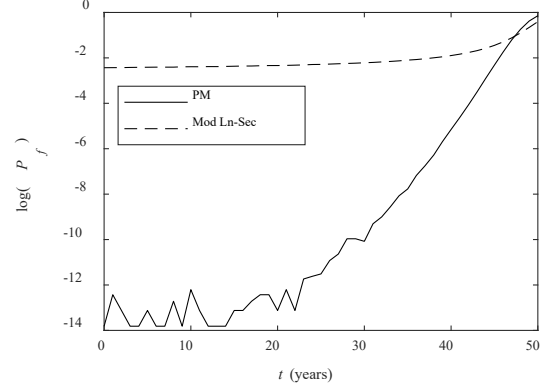
(a) Crack defect depth growth



(b) Failure prediction, P_b

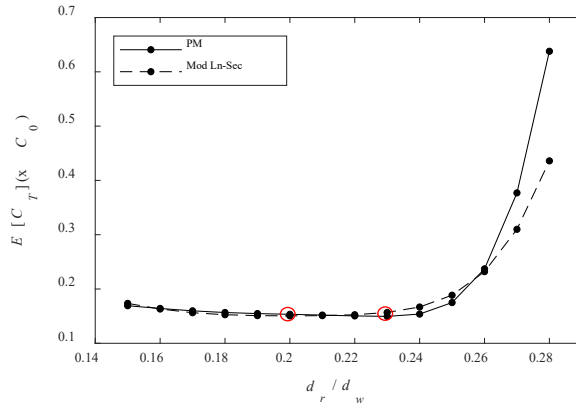


(c) Probability of failure, P_f

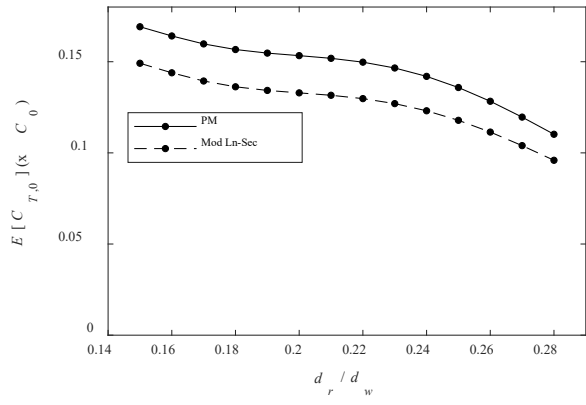


(d) Probability of failure, $\log(P_f)$

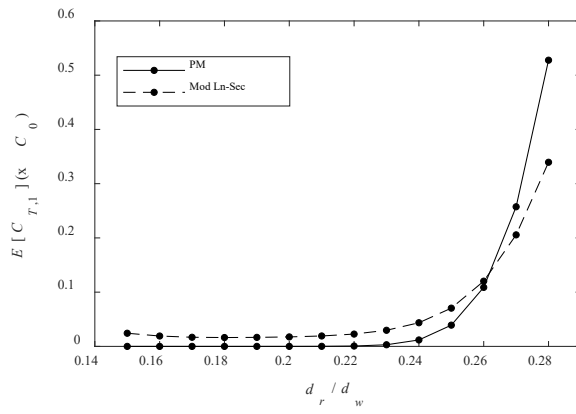
Figure 102. Time evolution of a/d_w , P_b , and P_f calculated for Case Study III



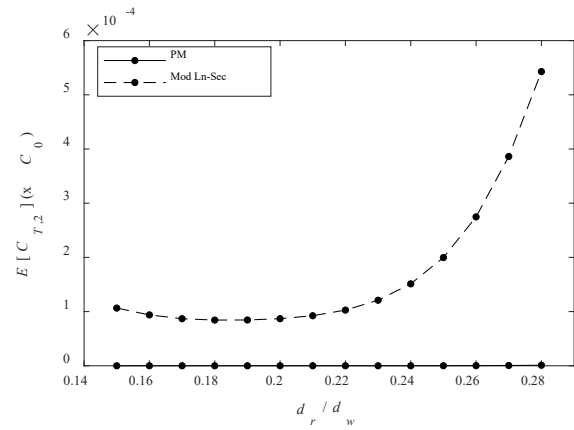
(a)



(b)



(c)



(d)

Figure 103. Expected total cost versus inspection interval considering $P_f(t)$ evaluated by the proposed model (PM) and the modified Ln-Sec for pipelines with single crack-like defect

7 Conclusions

To achieve the 1st objective, the ground low-carbon steel samples of flat sheet and pre-damaged sheet were exposed to two different environments (ASTM B117 and G85) to generate realistic corrosion profiles. The flat samples corroded during the exposure under two environments demonstrated by increased corrosion depth through IFM characterization; however, the sample variations were very big. The pre-damaged samples were sensitive to B117 environment that had increased corrosion depth of the scratched area with small sample variations; however, the depth change was less significant under the G85 environment, and the sample variations were large. The pre-damaged samples in B117 exposure were for NDE characterization.

To achieve the 2nd objective, the MSU NDE team has been tackling the challenge of interactive defects detection and characterization in metallic pipes, crucial for pipeline integrity assessment. The team developed a multi-modal electromagnetic and ultrasonic framework that leverages the benefits of various NDE and data processing methods such as machine learning. This framework includes Shear Horizontal (SH) guided wave testing, which has proven effective for NDE of buried pipelines, a task that typically presents significant field-testing challenges. In terms of Eddy Current Array (ECA) work, the team developed the ECA sensing probe, robust post-processing operations to convert raw eddy current data into clear 2D voltage mapped images of the defects. Regarding ultrasonic NDE work, the team focused on establishing an expanded NDE framework for interactive anomalies by probabilistically characterizing defect profiles. They used finite element modeling (FEM) for accurate defect modeling and to study the resulting ultrasonic NDE response. The team also developed numerical models that simulated different materials and structural conditions, to obtain their corresponding ultrasonic response for the complex anomaly scenario. MSU team also developed Machine Learning based feature engineering algorithms and Deep Learning based Convolutional Neural Networks (CNNs) for better characterizing the identified interactive anomalies. Despite high noise levels, the performance of the networks remained acceptable.

To achieve the 3rd objective, three tasks are performed. The first task is to develop probabilistic failure pressure models for pipelines with a single corrosion defect using a comprehensive database established in this study. The database is established by collecting data from literature and adding new numerical data generated from finite element models. With the established database, the performance of existing failure pressure prediction models for pipelines

with a single corrosion defect is compared. Then probabilistic failure pressure models are proposed using multivariate linear regression with existing failure pressure prediction models as independent variables. Also, a sensitive analysis is performed to evaluate the impact of influencing quantities on the proposed models.

The second task to meet the 3rd objective is to develop a probabilistic interaction rule and failure pressure prediction model for pipelines with interacting corrosion defects based on a comprehensive database established in this study. The database is established by collecting data from literature and adding new numerical data generated from finite element models. Using the database, the performance comparison of existing interaction rule is conducted first. Then a probabilistic interaction rule is proposed by using the logistic regression algorithm with pipe properties and adjacent defects characteristics (e.g., defect, length, and width of defects and spacing dimensions between defects) as independent variables. The proposed interaction rule is also compared with the existing interaction rules. Next, existing failure prediction models are compared using the established database. Then, the proposed failure pressure is developed by adding a corrosion factor to the best existing model identified. The correction factor is formulated using multivariate linear regression with pipe properties and adjacent defects characteristics as independent variables.

The third task to meet the 3rd objective is to develop a probabilistic failure pressure model is developed for pipelines with a single crack-like defect using a comprehensive database established in this study. The database is established by collecting data from literature and adding new numerical data generated from finite element models. In particular, extended finite element method (XFEM) is utilized to overcome the challenge of modeling cracking growth with the conventional finite element method, which is the need of mesh conformity to the geometry discontinuities and remeshing as crack grows. With the established database, a performance comparison of existing failure pressure prediction models for pipelines with a single crack-like defect is conducted. Then, the proposed failure pressure model is developed by adding a correction factor to the best existing model identified. The correction factor is modeled using a multivariate linear regression with pipe properties and crack defect characteristics as independent variables. Then, the proposed model is compared with the best existing model and a sensitive analysis is performed to evaluate the impact of influencing quantities on the proposed model.

To meet the 4th objective, three case studies are conducted. In the Case Study I, a reliability

analysis is performed to assess the impact of the failure pressure model of pipeline with a single corrosion defect on the structural performance. In the Case Study II, the time-dependent probability of the burst failure of a pipeline with two corrosion defects is evaluated, which is used to understand the importance of the prediction model for defect interaction identification in the pipeline integrity prediction. In Case Study III, a life-cycle cost analysis of a pipeline with a single crack-like defect is performed to evaluate the influence of the predicted failure pressure of pipeline with cracking on the expected total life cycle cost of pipelines.

Based on the results of this study, the following conclusions are drawn:

- The ground low-carbon steel can corrode under ASTM B117 and ASTM G85 exposure conditions. Pre-damaged samples showed increased corrosion depth and small sample variations during 36 weeks of exposure in B117. Large sample variations were found for the pre-damaged samples under G85 exposure and for flat samples.
- IFM is a powerful tool to characterize corrosion depth on the metal surface.
- The SH waves, having little out-of-plane leakage and thus being able to propagate over long distances within the pipe walls, are excellent for detecting local changes in thickness or material degradation.
- The ECA results indicated that damage could be clearly seen and monitored over time, showing the growth of corrosion and the deformation of the original defect. Despite some limitations related to saturation in later samples and some bias in the data, future improvements should include removing bias between coils and using precision tilt mechanisms for calibration.
- Ultrasonic GWs successfully propagated in a pipeline using a 2-D FEM based model and also corrosion pits and their ultrasonic NDE response are successfully modelled.
- Most existing burst failure pressure prediction models for pipelines with single corrosion defects are conservative, but the models that use the strain-hardening behavior of pipelines have good prediction performance. Particularly, model G5-19 developed by Zhu & Leis [10], is found to be the best existing model.
- Among the existing interaction rules considered, the POF, API RP 579, and 6WT interaction rules have better performance.
- Among the existing failure pressure prediction models for pipelines with interacted corrosion defects, Level-2 assessment methods such as the RSTRENG Effective Area

method, the DNV RP-F101 method for interacting defects (Part B), and the MTI method have better performance than Level-1 assessment methods; and the MTI method performs the best.

- Among the existing failure pressure prediction models for pipelines with crack-like defects (i.e., Ln-Sec, modified Ln-Sec, Corlas, API RP 579, and BS 7910), the modified Ln-Sec performs the best.
- FEM constructed in ABQUS is successfully used for burst testing simulation for pipelines with single corrosion defect or interacted corrosion defects; XFEM constructed in ABAQUS is successfully used for burst testing simulation for pipelines with crack-like defects.
- The sensitivity analysis reveals that the proposed models for pipelines with single corrosion defects and model G5-19 are sensitive to the pipe thickness and the depth and length of the corrosion defect. Also, both models are more sensitive to the defect length for deeper defects but become less sensitive for long defects length.
- The proposed interaction rule that is a function of colony configuration, defect geometries, pipe material and geometrical properties is more accurate than all the existing interaction rules used in this study.
- The correction factor proposed improves the MTI method for the failure pressure prediction of pipeline with interacted corrosion defects.
- The correction factor proposed improves the modified Ln-Sec by reducing the variability in the prediction, which could offer more accurate performance evaluation for risk management.
- The sensitivity study shows that the proposed pressure model for pipelines with crack-like defects is sensitive to yielding strength, cracking geometries, and the pipe all thickness ratio; in particular, the proposed model is more sensitive to the crack depth than the existing model, the modified Ln-Sec model.
- The reliability analysis in Case Study I shows that the depth and length of corrosion defect are crucial in the reliability evaluation of corroded pipelines, especially when the corrosion become worse; the performance of a failure pressure prediction model plays a critical role in determining the structural performance of corroded pipelines; and one should not ignore the uncertainty in the model error of the failure pressure prediction model.

- The importance analysis in Case Study I shows that wall thickness and model error in the burst pressure prediction model are the two most important random variables that contribute to the variability of the pipeline performance. This result shows that reducing the uncertainty in the wall thickness estimation is important and continuing improvement burst pressure prediction model can make significant impact.
- The time-dependent reliability curves obtained in Case Study II show that defect interaction significantly decreases the reliability of the pipeline, indicating capturing the interaction effect is critical in the corrosion risk management of pipelines.
- The expected life cycle costs obtained in Case Study III show that the burst pressure prediction model used to evaluate the probability of failure plays a role in the risk management of pipelines.

7.1 Net Safety Impact

The results of this study address DOT's pressing need to maintain safety and integrity of the existing pipeline infrastructure in the U.S under interactive threats. An expanded and new multi-modal NDE framework is developed enabling the missing capability to assess interactive anomalies with integration of lab-, field- and simulation-environment validation. A crucial body of knowledge of interactive anomalies and their properties has been established and facilitated for future design of assessment models and standards. In addition, the propagation of the anomaly profile (either isolated or colony defect) will be captured through probabilistic defect time-evaluation models based on NDE data. The improved accuracy in the pressure failure prediction and the appropriate consideration of defect interaction and prevailing uncertainties facilitate the development and application of quantitative risk management for pipelines. The use of reliable performance predictions (through the proposed NDE, defect time-evolution model, and failure pressure model) enables optimum monitoring/inspection, maintenance scheduling/methods, repair strategies/methods, and financial resource allocations and forecasting.

7.2 Future Work

The lab testing condition to generate cracking corrosion needs to be studied and understood in the future. The corrosion profile of the cracking corrosion can also be investigated if it can be easily generated in lab conditions.

Regarding NDE, research will focus on multi-modal data fusion that combines ECA, UT, and IFM data at both the measured data-level and feature-level. Time-dependent corrosion process

and damage characterization will be performed with uncertainties quantification and will be used for failure pressure prediction.

Concerning pipelines with crack-like defects, the probabilistic model was developed for thin-walled pipelines containing single crack-like defects. However, two other scenarios (i.e., colonies of crack-like defects or corrosion and crack-like defects) can occur on the surface of a pipeline. Therefore, a review of existing methodologies for determining the failure pressure for pipelines with interacting crack-like defects and existing rules for identifying cracking defect interaction need to be conducted. Due to extremely limited testing data available for these two scenarios, research is needed for conducting burst testing. With a reasonable database, improvement on the burst pressure models and interaction rules can be made for more accurate prediction.

References

- [1] Y. Gotoh, H. Hirano, M. Nakano, K. Fujiwara, and N. Takahashi, "Electromagnetic nondestructive testing of rust region in steel," *IEEE Trans Magn*, vol. 41, no. 10, pp. 3616–3618, 2005.
- [2] K. Gadonna, O. Leroy, P. Leprince, C. Boisse-Laporte, and L. L. Alves, "Study of gas heating by a microwave plasma torch," *Journal of Modern Physics*, vol. 3, no. 10, p. 1603, 2012.
- [3] Y. He, G. Tian, H. Zhang, M. Alamin, A. Simm, and P. Jackson, "Steel corrosion characterization using pulsed eddy current systems," *IEEE Sens J*, vol. 12, no. 6, pp. 2113–2120, 2012.
- [4] G. Van Drunen and V. S. Cecco, "Recognizing limitations in eddy-current testing," *NDT international*, vol. 17, no. 1, pp. 9–17, 1984.
- [5] ASME B31G, "Manual for determining the remaining strength of corroded pipelines," 2012.
- [6] D. Ritchie and S. Last, "Burst criteria of corroded pipelines – defect acceptance criteria," in *EPRG-PRCI 10th Biannual Joint Technical Meeting on Pipeline Research*, Cambridge, UK, 1995.
- [7] A. C. Benjamin and E. Q. de Andrade, "Modified method for the assessment of the remaining strength of corroded pipelines," in *Rio pipeline 2003 conference and exposition*, Rio de Janeiro, 2003.
- [8] DNV, "DNV-RP-F101: Recommended practice, Corroded pipelines," 2015. [Online]. Available: www.dnvgl.com.
- [9] T. A. Netto, U. S. Ferraz, and S. F. Estefen, "The effect of corrosion defects on the burst pressure of pipelines," *J Constr Steel Res*, vol. 61, no. 8, pp. 1185–1204, Aug. 2005, doi: 10.1016/j.jcsr.2005.02.010.
- [10] Z. Mustaffa and P. van Gelder, "A review and probabilistic analysis of limit state functions of corroded pipelines," in *Proceedings of the Twentieth International Offshore and Polar Engineering Conference*, Beijing, China: International Society of Offshore and Polar Engineers, 2010.
- [11] N. Wang and M. S. Zarghamee, "Evaluating Fitness-for-Service of Corroded Metal Pipelines: Structural Reliability Bases," *J Pipeline Syst Eng Pract*, vol. 5, no. 1, Feb. 2014, doi: 10.1061/(ASCE)PS.1949-1204.0000148.
- [12] X.-K. Zhu and B. N. Leis, "Influence of Yield-to-Tensile Strength Ratio on Failure Assessment of Corroded Pipelines," *J Press Vessel Technol*, vol. 127, no. 4, Nov. 2005, doi: 10.1115/1.2042481.
- [13] R. Amaya-Gómez, M. Sánchez-Silva, E. Bastidas-Arteaga, F. Schoefs, and F. Muñoz, "Reliability assessments of corroded pipelines based on internal pressure – A review," *Engineering Failure Analysis*, vol. 98, Elsevier Ltd, pp. 190–214, Apr. 01, 2019. doi: 10.1016/j.engfailanal.2019.01.064.
- [14] Z. Chen, Y. Chen, W. Wang, K. Lu, H. Yang, and W. Zhu, "Failure pressure analysis of hydrogen storage pipeline under low temperature and high pressure," *Int J Hydrogen Energy*, vol. 45, no. 43, pp. 23142–23150, Sep. 2020, doi: 10.1016/j.ijhydene.2020.06.129.

- [15] W. Zhou and G. X. Huang, "Model error assessments of burst capacity models for corroded pipelines," *International Journal of Pressure Vessels and Piping*, vol. 99–100, pp. 1–8, Nov. 2012, doi: 10.1016/j.ijpvp.2012.06.001.
- [16] D. S. Cronin and R. J. Pick, "Prediction of the failure pressure for complex corrosion defects," *International Journal of Pressure Vessels and Piping*, vol. 79, no. 4, Apr. 2002, doi: 10.1016/S0308-0161(02)00020-0.
- [17] CSA, "CSA Z662-07: Limit state equation for burst of large leaks and rupture for corrosion defect," 2007.
- [18] D. R. Stephens and B. N. Leis, "Development of an Alternative Criterion for Residual Strength of Corrosion Defects in Moderate- to High-Toughness Pipe," in *Volume 2: Integrity and Corrosion; Offshore Issues; Pipeline Automation and Measurement; Rotating Equipment*, American Society of Mechanical Engineers, Oct. 2000. doi: 10.1115/IPC2000-192.
- [19] S. Hasan, F. Khan, and S. Kenny, "Probability assessment of burst limit state due to internal corrosion," *International Journal of Pressure Vessels and Piping*, vol. 89, pp. 48–58, Jan. 2012, doi: 10.1016/j.ijpvp.2011.09.005.
- [20] B. Flu and A. D. Batte, "An overview of advanced methods for the assessment of corrosion in linepipe," 1999.
- [21] J. F. Kiefner and A. R. Duffy, "Summary of research to determine the strength of corroded areas in line pipe, Technical Report, Battelle Columbus Laboratories, 1971.,," 1971.
- [22] Y. Chen, H. Zhang, J. Zhang, X. Li, and J. Zhou, "Failure analysis of high strength pipeline with single and multiple corrosions," *Mater Des*, vol. 67, Feb. 2015, doi: 10.1016/j.matdes.2014.10.088.
- [23] J. B. Choi, B. K. Goo, J. C. Kim, Y. J. Kim, and W. S. Kim, "Development of limit load solutions for corroded gas pipelines," *International Journal of Pressure Vessels and Piping*, vol. 80, no. 2, pp. 121–128, Apr. 2003, doi: 10.1016/S0308-0161(03)00005-X.
- [24] W. Zhou and G. X. Huang, "Model error assessments of burst capacity models for corroded pipelines," *International Journal of Pressure Vessels and Piping*, vol. 99–100, pp. 1–8, Nov. 2012, doi: 10.1016/j.ijpvp.2012.06.001.
- [25] S. Sheather, *A Modern Approach to Regression with R*. Springer Science & Business Media, 2009.
- [26] J. F. Kiefner and P. H. Vieth, "PC program speeds new criterion for evaluating corroded pipe," *Oil Gas Journal*, vol. 88, no. 32, pp. 91–93, Aug. 1990.
- [27] Pipeline Operators Forum, "Specifications and requirements for intelligent pig inspection of pipelines," 2009.
- [28] API 579-1/ASME FFS-1, "Fitness for service," *American Society of Mechanical Engineers*. USA, 2016.
- [29] BS 7910, "Guide to methods for assessing the acceptability of flaws in metallic structures," *British Standards Institute*. UK, 2013.

- [30] M. Lamontagne, "Interaction Rules - An Integral Factor," in *NACE international*, Denver, Apr. 2002.
- [31] X. Li, Y. Bai, C. Su, and M. Li, "Effect of interaction between corrosion defects on failure pressure of thin wall steel pipeline," *International Journal of Pressure Vessels and Piping*, vol. 138, pp. 8–18, Feb. 2016, doi: 10.1016/j.ijpvp.2016.01.002.
- [32] B. C. Mondal and A. S. Dhar, "Interaction of multiple corrosion defects on burst pressure of pipelines," *Canadian Journal of Civil Engineering*, vol. 44, no. 8, pp. 589–597, Aug. 2017, doi: 10.1139/cjce-2016-0602.
- [33] H. Zhang and Z. Tian, "An Integrated Reliability Method with a Newly Developed Interaction Rule for Steel Pipelines with Multiple Corrosion Defects," *J Pipeline Syst Eng Pract*, vol. 13, no. 4, Nov. 2022, doi: 10.1061/(ASCE)PS.1949-1204.0000683.
- [34] A. C. Benjamin, E. Q. de Andrade, B. P. Jacob, L. C. Pereira, and P. R. S. Machado, "Failure Behavior of Colonies of Corrosion Defects Composed of Symmetrically Arranged Defects," in *Volume 2: Integrity Management; Poster Session; Student Paper Competition*, ASMEDC, Jan. 2006, pp. 417–432. doi: 10.1115/IPC2006-10266.
- [35] A. C. Benjamin, J. L. F. Freire, and R. D. Vieira, "Analysis of pipeline containing interacting corrosion defects," *Exp Tech*, vol. 31, no. 3, pp. 74–82, May 2007, doi: 10.1111/j.1747-1567.2007.00190.x.
- [36] A. C. Benjamin, J. L. F. Freire, R. D. Vieira, and D. J. S. Cunha, "Interaction of corrosion defects in pipelines – Part 1: Fundamentals," *International Journal of Pressure Vessels and Piping*, vol. 144, pp. 56–62, Aug. 2016, doi: 10.1016/j.ijpvp.2016.05.007.
- [37] Y. Chen, H. Zhang, J. Zhang, X. Li, and J. Zhou, "Failure analysis of high strength pipeline with single and multiple corrosions," *Mater Des*, vol. 67, pp. 552–557, Feb. 2015, doi: 10.1016/j.matdes.2014.10.088.
- [38] A. C. Benjamin, J. L. F. Freire, R. D. Vieira, and E. Q. de Andrade, "Burst tests on pipeline containing closely spaced corrosion defects," in *Volume 4: Terry Jones Pipeline Technology; Ocean Space Utilization; CFD and VIV Symposium*, ASMEDC, Jan. 2006, pp. 103–116. doi: 10.1115/OMAE2006-92131.
- [39] B. Flu and A. D. Batte, "An overview of advanced methods for the assessment of corrosion in linepipe," 1999.
- [40] E. Q. de Andrade *et al.*, "Finite element modeling of the failure behavior of pipelines containing interacting corrosion defects," in *Volume 4: Terry Jones Pipeline Technology; Ocean Space Utilization; CFD and VIV Symposium*, ASMEDC, Jan. 2006, pp. 315–325. doi: 10.1115/OMAE2006-92600.
- [41] A. C. Benjamin, J. L. F. Freire, R. D. Vieira, and E. Q. de Andrade, "Burst tests on pipeline containing closely spaced corrosion defects," in *Volume 4: Terry Jones Pipeline Technology; Ocean Space Utilization; CFD and VIV Symposium*, ASMEDC, Jan. 2006, pp. 103–116. doi: 10.1115/OMAE2006-92131.

- [42] K. J. Kere and Q. Huang, "Development of probabilistic failure pressure models for pipelines with single corrosion defect," *International Journal of Pressure Vessels and Piping*, vol. 197, p. 104656, Jun. 2022, doi: 10.1016/j.ijpvp.2022.104656.
- [43] A. Cosham, P. Hopkins, and B. Leis, "Crack-Like defects in pipelines: The relevance of pipeline-specific methods and standards," in *Volume 2: Pipeline Integrity Management*, American Society of Mechanical Engineers, Sep. 2012, pp. 713–726. doi: 10.1115/IPC2012-90459.
- [44] J. F. Kiefner, W. A. Maxey, R. J. Eiber, and A. R. Duffy, "Failure stress levels of flaws in pressurized cylinders," in *Progress in Flaw Growth and Fracture Toughness Testing*, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428-2959: ASTM International, 1973, pp. 461–481. doi: 10.1520/STP49657S.
- [45] J. F. Kiefner, "Modified Ln-Secant equation improves failure prediction," *Oil & Gas Journal*, vol. 106, no. 38, Oct. 2008.
- [46] S. J. Polasik, C. E. Jaske, and T. A. Bubenik, "Review of engineering fracture mechanics model for pipeline applications," in *Volume 1: Pipelines and Facilities Integrity*, American Society of Mechanical Engineers, Sep. 2016. doi: 10.1115/IPC2016-64605.
- [47] M. Staat, "Plastic collapse analysis of longitudinally flawed pipes and vessels," *Nuclear Engineering and Design*, vol. 234, no. 1–3, pp. 25–43, Dec. 2004, doi: 10.1016/j.nucengdes.2004.08.002.
- [48] T. L. Anderson, "Development of a modern assessment method for longitudinal seam weld cracks," *PRCI Catalogue No. PR-460-134506-R0*. Pipeline Research Council International, Chantilly, 2015.
- [49] S. Tandon, M. Gao, R. Krishnamurthy, S. Kariyawasam, and R. Kania, "Evaluation of existing fracture mechanics models for burst pressure predictions, pheoretical and pxperimental aspects," in *Volume 2: Pipeline Integrity Management*, American Society of Mechanical Engineers, Sep. 2014. doi: 10.1115/IPC2014-33563.
- [50] S. A. Hosseini, "Assessment of Crack in Corrosion Defects in Natural Gas Transmission Pipelines," University of Waterloo, Waterloo, 2010.
- [51] Z. Yan, S. Zhang, and W. Zhou, "Model error assessment of burst capacity models for energy pipelines containing surface cracks," *International Journal of Pressure Vessels and Piping*, vol. 120–121, pp. 80–92, Aug. 2014, doi: 10.1016/j.ijpvp.2014.05.007.
- [52] EDF, "R6 : Assessment of the integrity of the structures containing defects, Revision 4, Amendement 10." EDF Energy, Gloucester, 2013.
- [53] Simulia, "Abaqus analysis user's guide: Modeling discontinuities as an enriched feature using the extended finite element method."
- [54] T. Belytschko and T. Black, "Elastic crack growth in finite elements with minimal remeshing," *Int J Numer Methods Eng*, vol. 45, no. 5, pp. 601–620, Jun. 1999, doi: 10.1002/(SICI)1097-0207(19990620)45:5<601::AID-NME598>3.0.CO;2-S.

- [55] M. Lin, S. Agbo, J. J. R. Cheng, N. Yoosef-Ghodsi, and S. Adeeb, "Application of the extended finite element method (XFEM) to simulate crack propagation in pressurized steel pipes," in *Volume 3B: Design and Analysis*, American Society of Mechanical Engineers, Jul. 2017. doi: 10.1115/PVP2017-65575.
- [56] A. Okodi, M. Lin, N. Yoosef-Ghodsi, M. Kainat, S. Hassanien, and S. Adeeb, "Crack propagation and burst pressure of longitudinally cracked pipelines using extended finite element method," *International Journal of Pressure Vessels and Piping*, vol. 184, p. 104115, Jul. 2020, doi: 10.1016/j.ijpvp.2020.104115.
- [57] T. L. Anderson, *Fracture mechanics: Fundamentals and applications*, 3rd ed. Boca Raton, Florida: CRC Press, 2005.
- [58] M. D. Rana and G. B. Rawls, "Prediction of fracture stresses of high pressure gas cylinders containing cracklike flaws," *J Press Vessel Technol*, vol. 129, no. 4, pp. 639–643, Nov. 2007, doi: 10.1115/1.2767346.
- [59] A. Okodi, M. Lin, N. Yoosef-Ghodsi, M. Kainat, S. Hassanien, and S. Adeeb, "Crack propagation and burst pressure of longitudinally cracked pipelines using extended finite element method," *International Journal of Pressure Vessels and Piping*, vol. 184, p. 104115, Jul. 2020, doi: 10.1016/j.ijpvp.2020.104115.
- [60] S. A. Miran, Q. Huang, and H. Castaneda, "Time-Dependent Reliability Analysis of Corroded Buried Pipelines Considering External Defects," *Journal of Infrastructure Systems*, vol. 22, no. 3, Sep. 2016, doi: 10.1061/(ASCE)IS.1943-555X.0000307.
- [61] O. Ditlevsen and H. O. Madsen, *Structural Reliability Methods*. Chichester: John Wiley & Sons Ltd, 1996.
- [62] A. D. Kiureghian and J. B. Ke, "Finite-element based reliability analysis of frame structure," *Conference on Structural Safety and Reliability*, pp. 395–404, 1985.
- [63] F. Caleyó, J. C. Velázquez, A. Valor, and J. M. Hallen, "Probability distribution of pitting corrosion depth and rate in underground pipelines: A Monte Carlo study," *Corros Sci*, vol. 51, no. 9, pp. 1925–1934, Sep. 2009, doi: 10.1016/j.corsci.2009.05.019.
- [64] K. J. Kere and Q. Huang, "Risk management strategies and probabilistic failure pressure model development for pipelines with crack-like defect," in *14th International Conference on Applications of Statistics and Probability in Civil Engineering*, Dublin, Ireland: ICASP14, Jul. 2023.
- [65] P. Paris and F. Erdogan, "A critical analysis of crack propagation laws," *Journal of Basic Engineering*, vol. 85, no. 4, pp. 528–533, Dec. 1963, doi: 10.1115/1.3656900.
- [66] G. R. Irwin, "The crack-extension-force for a crack at a free surface boundary," Washington, DC, 1958.
- [67] W. J. S. Gomes and A. T. Beck, "Optimal inspection and design of onshore pipelines under external corrosion process," *Structural Safety*, vol. 47, pp. 48–58, Mar. 2014, doi: 10.1016/j.strusafe.2013.11.001.

- [68] S. M. Soliman and D. M. Frangopol, "Life-cycle management of fatigue-sensitive structures integrating inspection information," *Journal of Infrastructure Systems*, vol. 20, no. 2, Jun. 2014, doi: 10.1061/(ASCE)IS.1943-555X.0000169.
- [69] S. Cicero, R. Lacalle, R. Cicero, and D. Ferreño, "Assessment of local thin areas in a marine pipeline by using the FITNET FFS corrosion module," *International Journal of Pressure Vessels and Piping*, vol. 86, no. 5, May 2009, doi: 10.1016/j.ijpvp.2008.11.021.
- [70] H. C. Phan, A. S. Dhar, and B. C. Mondal, "Revisiting burst pressure models for corroded pipelines," *Canadian Journal of Civil Engineering*, vol. 44, no. 7, pp. 485–494, 2017, doi: 10.1139/cjce-2016-0519.
- [71] T. A. Netto, U. S. Ferraz, and S. F. Estefen, "The effect of corrosion defects on the burst pressure of pipelines," *J Constr Steel Res*, vol. 61, no. 8, pp. 1185–1204, Aug. 2005, doi: 10.1016/j.jcsr.2005.02.010.
- [72] T. A. Netto, "A simple procedure for the prediction of the collapse pressure of pipelines with narrow and long corrosion defects - Correlation with new experimental data," *Applied Ocean Research*, vol. 32, no. 1, pp. 132–134, Feb. 2010, doi: 10.1016/j.apor.2009.12.007.
- [73] K. J. Yeom, Y. K. Lee, K. H. Oh, and W. S. Kim, "Integrity assessment of a corroded API X70 pipe with a single defect by burst pressure analysis," *Eng Fail Anal*, vol. 57, pp. 553–561, Nov. 2015, doi: 10.1016/j.engfailanal.2015.07.024.
- [74] R. Bea, R. Iversen, A. Kareem, and X. Chen, "Risk Assessment and Management (RAM) Based Guidelines for Requalification of Marine Pipelines," 1999.
- [75] X.-K. Zhu, "A New Material Failure Criterion for Numerical Simulation of Burst Pressure of Corrosion Defects in Pipelines," in *Volume 6B: Materials and Fabrication*, American Society of Mechanical Engineers, Jul. 2015. doi: 10.1115/PVP2015-45713.
- [76] Y. Shuai, J. Shuai, and K. Xu, "Probabilistic analysis of corroded pipelines based on a new failure pressure model," *Eng Fail Anal*, vol. 81, pp. 216–233, Nov. 2017, doi: 10.1016/j.engfailanal.2017.06.050.

Appendix A: Exposure Testing Results of Task 1
Surface photos of flat samples after exposure testing



SAMPLE 1



SAMPLE 2

Figure A1. The flat testing samples after 24 weeks of exposure in B117.



SAMPLE 1



SAMPLE 2

Figure A2. The flat testing samples after 28 weeks of exposure in G85 testing.

Pre-damaged samples exposure testing and IFM analysis

Exposure testing

The upper and lower sides around the scratch were covered by black tape to prevent corrosion, which is the black region in Figure B3. Then, the metal samples were placed in a salt fog chamber for exposure testing. The samples were continuously sprayed with 5 wt.% NaCl solution circulating at 35° C (ASTM B117 standard condition). ASTM G85 standard where samples were in exposure to a mixed solution consisting of 0.35 wt.% aluminum sulfate and 0.05 wt.% NaCl.

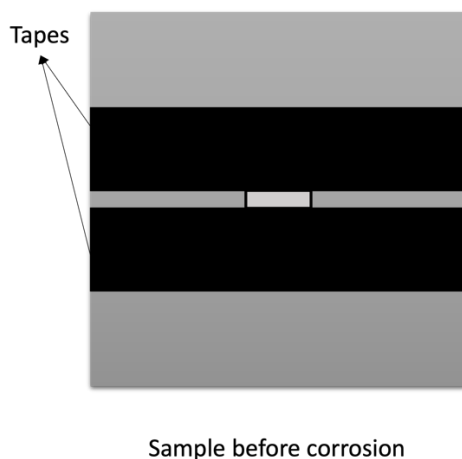


Figure A3. Schematic of a metal sample for exposure testing.

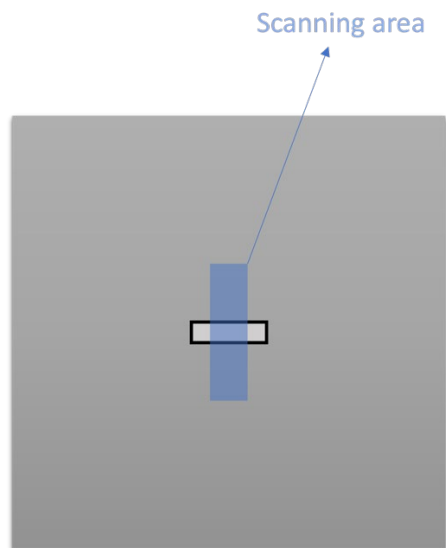
Surface cleaning

After salt fog exposure testing, the tapes were removed. The punched scratch region was washed with a mixed solution of HCl, Sb₂O₃, and SnCl₂ to remove corrosion products.

IFM characterization

The depth of the exposed metal samples was observed by IFM. The blue region in Figure B4 is the IFM scanning area. It covered the scratch and the tape-protected region of the metal. The protected region was used as a reference point for the measurement. The first scan of the blue area gave a depth distribution, and it also presented the lowest point in the punched scratch region. Then, a second measurement used a line profile that went through the lowest point and the reference point to measure the depth. An example of the line profile is shown in Figure B5.

The depth before exposure testing was measured exactly the same approach by IFM. The scanning area didn't cover the left and right end of the punched scratch because these areas were sloping surfaces of the corners which were not smooth.



Sample after corrosion and washing

Figure A4. Schematic of IFM scanning area for a metal sample.

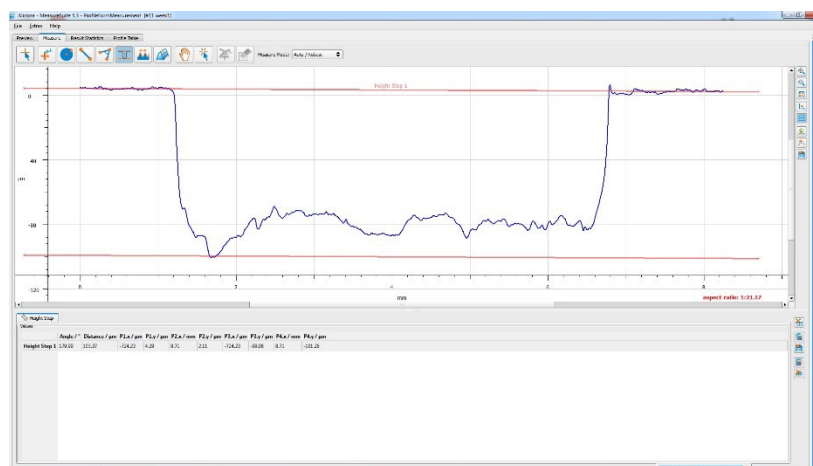


Figure A5. Line profile of a metal sample (#11 after 1-week exposure).

Surface photos of pre-damaged samples



Figure A6. The pre-damaged steel sample before the exposure testing.

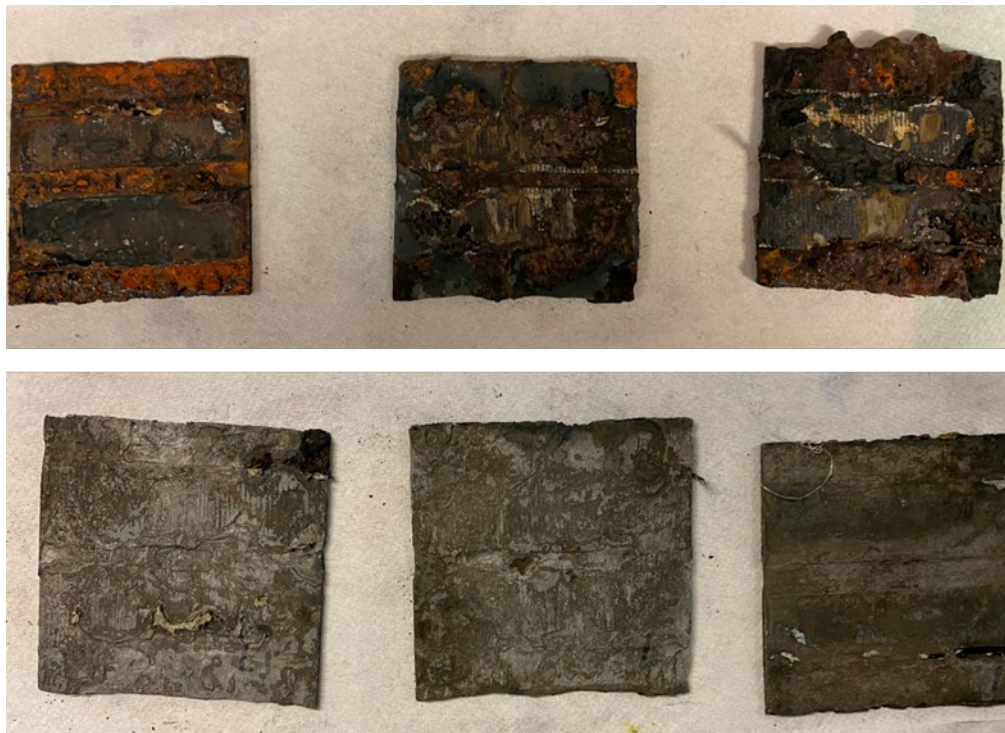


Figure A7. The pre-damaged testing samples after 24 weeks of exposure testing in B117: before surface cleaning (top) and after surface cleaning (bottom).

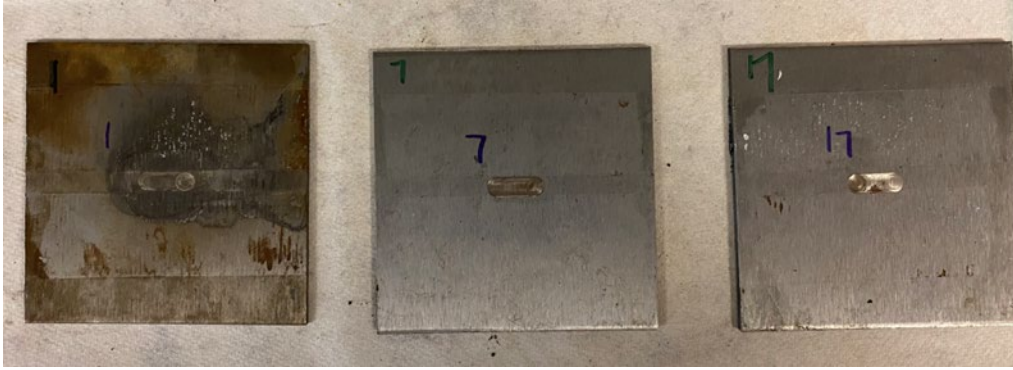


Figure A8. The pre-damaged testing samples after 24 weeks of exposure testing in G85.

Appendix B: Existing models of burst failure prediction for pipeline with single corrosion defect

G1-1 - ASME B31G Original [3]

$$\hat{y}_1 = \begin{cases} \frac{2t}{D} (1.1\sigma_y) \left[\frac{1 - (2/3)(d/t)}{1 - (2/3)(d/t)M^{-1}} \right], & l^2/Dt \leq 20 \\ \frac{2t}{D} (1.1\sigma_y) [1 - (d/t)] & , \quad l^2/Dt > 20 \end{cases} \quad (\text{B.1})$$

$$M = \sqrt{1 + 0.8 \left(\frac{l^2}{Dt} \right)} \quad (\text{B.1.a})$$

G1-2 - Modified B31G (RSTRENG 0.85dL) [3]

$$\hat{y}_2 = \frac{2t}{D} (\sigma_y + 69[\text{MPa}]) \left[\frac{1 - 0.85(d/t)}{1 - 0.85(d/t)M^{-1}} \right] \quad (\text{B.2})$$

$$M = \begin{cases} \sqrt{1 + 0.6275 \left(\frac{l^2}{Dt} \right) - 0.003375 \left(\frac{l^2}{Dt} \right)^2}, & l^2/Dt \leq 50 \\ 3.3 + 0.032 \left(\frac{l^2}{Dt} \right) & , \quad l^2/Dt > 50 \end{cases} \quad (\text{B.2.a})$$

G1-3 - SHELL92 [4]

$$\hat{y}_3 = \frac{2t(0.9\sigma_u)}{D} \left[\frac{1 - d/t}{1 - (d/t)M^{-1}} \right] \quad (\text{B.3})$$

$$M = \sqrt{1 + 0.8 \left(\frac{l^2}{Dt} \right)} \quad (\text{B.3.a})$$

G1-4 - RPA (Rectangular Parabolic Area) [5]

$$\hat{y}_4 = \frac{2t}{D} (\sigma_y + 69[\text{MPa}]) \left[\frac{1 - a(d/t)}{1 - a(d/t)M^{-1}} \right] \quad (\text{B.4})$$

$$a = \begin{cases} 0.85 & , \quad l^2/Dt \leq 20 \\ 1 - \frac{0.15(64 \times 10^6)}{\left(\frac{l^2}{Dt} \right)^6} & , \quad l^2/Dt > 20 \end{cases} \quad (\text{B.4.a})$$

$$M = \begin{cases} \sqrt{1 + 0.6275 \left(\frac{l^2}{Dt} \right) - 0.003375 \left(\frac{l^2}{Dt} \right)^2}, & l^2/Dt \leq 20 \\ 2.1 + 0.07 \left(\frac{l^2}{Dt} \right) & , \quad l^2/Dt > 20 \end{cases} \quad (\text{B.4.b})$$

GI-5 - RSTRENG Effective Area [3]

$$\hat{y}_5 = \frac{2t}{D} (\sigma_y + 69[MPa]) \left[\frac{1 - A/A_0}{1 - (A/A_0)M^{-1}} \right] \quad (B.5)$$

$$M = \begin{cases} \sqrt{1 + 0.6275 \left(\frac{l^2}{Dt} \right) - 0.003375 \left(\frac{l^2}{Dt} \right)^2}, & l^2/Dt \leq 50 \\ 3.3 + 0.032 \left(\frac{l^2}{Dt} \right), & l^2/Dt > 50 \end{cases} \quad (B.5.a)$$

GI-6 - CSA Z662 [14]

$$\hat{y}_6 = \begin{cases} P + 0.00228\sigma_u, & \sigma_y > 241MPa \\ 1.34P - 0.34P_0 + 0.00938\sigma_y, & \sigma_y \leq 241MPa \end{cases} \quad (B.6)$$

$$P = P_0 \left[\frac{1 - (d/t)}{1 - (d/t)M^{-1}} \right] \quad (B.6.a)$$

$$P_0 = \begin{cases} 1.8 \frac{t\sigma_u}{D}, & \sigma_y > 241MPa \\ 2.3 \frac{t\sigma_y}{D}, & \sigma_y \leq 241MPa \end{cases} \quad (B.6.b)$$

$$M = \begin{cases} \sqrt{1 + 0.6275 \left(\frac{l^2}{Dt} \right) - 0.003375 \left(\frac{l^2}{Dt} \right)^2}, & l^2/Dt \leq 50 \\ 3.3 + 0.032 \left(\frac{l^2}{Dt} \right), & l^2/Dt > 50 \end{cases} \quad (B.6.c)$$

GI-7 - DNV RP-F101 Method for single defects (Part B) [6]

$$\hat{y}_7 = \frac{2t\sigma_u(1 - (d/t))}{(D - t) \left(1 - \frac{(d/t)}{M} \right)} \quad (B.7)$$

$$M = \sqrt{1 + 0.31 \left(\frac{l^2}{Dt} \right)} \quad (B.7.a)$$

GI-8 - Fitnet FFS [18]

$$\hat{y}_8 = \frac{2t\sigma_u \left(\frac{1}{2} \right)^{65\sigma_y^{-1}}}{D - t} \left[\frac{1 - d/t}{1 - (d/t)M^{-1}} \right] \quad (B.8)$$

$$M = \sqrt{1 + 0.8 \left(\frac{l^2}{Dt} \right)} \quad (B.8.a)$$

G1-9: Phan et al. - Modified NG-18 [19]

$$\hat{y}_9 = \frac{2t\sigma_u}{D} \left[\frac{1 - 0.92126 \left(\frac{d}{t}\right)}{1 - 0.92126 \left(\frac{d}{t}\right) M^{-1}} \right] \quad (\text{B.9})$$

$$M = \left(1 + 0.06361 \left(\frac{l^2}{Dt} \right)^{2.75485} \right) \quad (\text{B.9.a})$$

G2-10: Netto et al. [7]

$$\hat{y}_{10} = \frac{2t(1.1\sigma_y)}{D} \left[1 - 0.9435 \left(\frac{d}{t}\right)^{1.6} \left(\frac{l}{D}\right)^{0.4} \right] \quad (\text{B.10})$$

G2-11: Mustaffa & van Gelder [8]

$$\hat{y}_{11} = \sigma_u \left(\frac{t}{D}\right)^{0.8442} \left(\frac{d}{t}\right)^{-0.0545} \left(\frac{l}{w}\right)^{-0.0104} \quad (\text{B.11})$$

G2-12: Netto et al. [20]

$$\hat{y}_{12} = \frac{2t(1.1\sigma_y)}{D} \left[\frac{1 - d/t}{1 - d/t \left(1 - \left(\frac{w}{\pi D}\right)^{0.4} \left(\frac{l}{10D}\right)^{0.4} \right)} \right]^{2.675} \quad (\text{B.12})$$

G2-13: Wang & Zarghamee [9]

$$\hat{y}_{13} = \begin{cases} P_{bi} \left[1 - 0.886 \left(\frac{d}{t}\right) \left(\frac{l}{D}\right)^{0.3} \right] & , \quad D < 610\text{mm} \\ P_{bi} \left[1 - 1.120 \left(\frac{d}{t}\right)^{1.15} \left(\frac{l}{D}\right)^{0.3} \right] & , \quad D \geq 610\text{mm} \end{cases} \quad (\text{B.13})$$

$$P_{bi} = \frac{2t(1.1\sigma_y)}{D} \quad (\text{B.13.a})$$

G2-14: Phan et al. - Modified Netto et al. [19]

$$\hat{y}_{14} = \frac{2t\sigma_u}{D} \left[1 - 0.88555 \left(\frac{d}{t}\right)^{0.98077} \left(\frac{l}{D}\right)^{0.31053} \right] \quad (\text{B.14})$$

G3-15: PCORRC (Pipe Corrosion Criterion) [15]

$$\hat{y}_{15} = \frac{2t\sigma_u}{D} \left[1 - \frac{d}{t} \left(1 - \exp \left\{ -0.157 \frac{l}{\sqrt{r(t-d)}} \right\} \right) \right] \quad (\text{B.15})$$

G3-16: Modified PCORRC [73]

$$\hat{y}_{16} = \frac{2t(0.9\sigma_u)}{D} \left[1 - \frac{d}{t} \left(1 - \exp \left\{ -0.224 \frac{l}{\sqrt{r(t-d)}} \right\} \right) \right] \quad (\text{B.16})$$

G4-17: Original RAM Pipe Requal [22]

$$\hat{y}_{17} = 2.4 \frac{t\sigma_u}{(D-t)\varphi} \quad (\text{B.17})$$

$$\varphi = 1 + 2 \sqrt{\frac{d}{r}} \quad (\text{B.17.a})$$

G4-18: Modified Ram Pipe Requal [22]

$$\hat{y}_{18} = 2.2(t-d) \frac{\sigma_u}{(D-t)\varphi} \quad (\text{B.18})$$

$$\varphi = 1 + 2 \sqrt{\frac{d}{r}} \quad (\text{B.18.a})$$

G5-19: Zhu & Leis [10]

$$\hat{y}_{19} = \frac{4}{(3)^{\frac{n+1}{2}}} \frac{t}{D} \sigma_u \left[1 - d/t \left(1 - \exp \left\{ -0.157 \frac{l}{\sqrt{R(t-d)}} \right\} \right) \right] \quad (\text{B.19})$$

$$n = -0.0319 + \sqrt{0.0856 \left(\frac{\sigma_u}{\sigma_y} \right) - 0.0846} \quad (\text{B.19.a})$$

G5-20: Zhu - X65 [23]

$$\hat{y}_{20} = \left(\frac{2 + \sqrt{3}}{4\sqrt{3}} \right)^{n+1} \frac{4t}{D} \sigma_u \left[1 - \frac{d}{t} f_L \right] \quad (\text{B.20})$$

$$f_L = 1 - \frac{1}{1 + 0.1385 \left(\frac{l}{\sqrt{Dt}} \right) + 0.1357 \left(\frac{l^2}{Dt} \right)} \quad (\text{B.20.a})$$

$$n = -0.0319 + \sqrt{0.0856 \left(\frac{\sigma_u}{\sigma_y} \right) - 0.0846} \quad (\text{B.20.b})$$

G6-21: Choi et al. [24]

$$\hat{y}_{21} = \begin{cases} \frac{2t}{D} (0.9\sigma_u) \left[C_0 + C_1 \left(\frac{l}{\sqrt{rt}} \right) + C_2 \left(\frac{l}{\sqrt{rt}} \right)^2 \right], & l/\sqrt{rt} < 6 \\ \frac{2t}{D} \sigma_u \left[C_3 + C_4 \left(\frac{l}{\sqrt{rt}} \right) \right] & , \quad l/\sqrt{rt} \geq 6 \end{cases} \quad (\text{B.21})$$

G6-22: Chen et al. [25]

$$\hat{y}_{22} = \frac{2t\sigma_u}{D-t} \mathcal{C} \quad (\text{B.22})$$

$$\mathcal{C} = \begin{cases} \left[C_0 \left(\frac{l}{\sqrt{Dt}} \right)^2 + C_1 \left(\frac{l}{\sqrt{Dt}} \right) + C_2 \right] \left[C_3 \left(\frac{w}{\pi D} \right)^2 + C_4 \left(\frac{w}{\pi D} \right) + C_5 \right], & l/\sqrt{Dt} \leq 5, w/\pi D \leq 0.3 \\ \left[C_6 \left(\frac{l}{\sqrt{Dt}} \right) + C_7 \right] \left[C_3 \left(\frac{w}{\pi D} \right)^2 + C_4 \left(\frac{w}{\pi D} \right) + C_5 \right] & , l/\sqrt{Dt} > 5, w/\pi D \leq 0.3 \\ \left[C_8 \left(\frac{l}{\sqrt{Dt}} \right)^2 + C_9 \left(\frac{l}{\sqrt{Dt}} \right) + C_{10} \right] & , l/\sqrt{Dt} \leq 5, w/\pi D > 0.3 \\ \left[C_{11} \left(\frac{l}{\sqrt{Dt}} \right) + C_{12} \right] & , l/\sqrt{Dt} > 5, w/\pi D > 0.3 \end{cases} \quad (\text{B.22.a})$$

G6-23: CUP [26]

$$\hat{y}_{23} = \frac{2t\sigma_u}{D} \left[1 - \frac{d}{t} \left(1 - \left(C_1 \left(1 - \left(\frac{w}{\pi D} \right)^2 \right)^6 + (1 - C_1) \exp \left\{ \frac{C_2 l}{\sqrt{Dt}} \right\} \right) \left(1 - \frac{d}{t} \right)^{C_3} \right) \right] \quad (\text{B.23})$$

where $C_1 = 0.1075$, $C_2 = -0.4103$, $C_3 = 0.2504$

G6-24: Phan et al. - Modified Gajdoš et al. [19]

$$\hat{y}_{24} = \frac{2t\sigma_u}{D} \left[1 - \frac{1.24678 \left(\frac{d}{t} \right)}{1 + 12.6739 \left(\frac{t}{l} \right)} \right] \quad (\text{B.24})$$