the Energy to Lead

FINAL REPORT
DOT Project No.: 643
Contract Number: DTPH5615T00007

Slow Crack Growth Evaluation of Vintage Polyethylene Pipes

Reporting Period:
October 1, 2015 through September 24, 2017

Report Issued:
October 3rd, 2017

Prepared for:
U.S. Department of Transportation Pipeline and Hazardous Materials Safety Administration
James Merritt
Technical Manager
Office of Pipeline Safety/Southwest Region
303-683-3117
James.Merritt@dot.gov

GTI Project Manager & Technical Contact:
Ernest Lever
Principal Investigator
Gas Technology Institute
ernest.lever@gastechnology.org
847-544-3415

GTI Team Members:
Saurav Acharya, Tishun Peng

Subcontractors:
Arizona State University
Michigan State University

Gas Technology Institute
1700 S. Mount Prospect Rd.
Des Plaines, Illinois 60018
www.gastechnology.org

File Name: DTPH5615T00007_Prj643_Final.docx

Signature Page

Print or typed First M. Last	Signature	Date
AUTHOR: Ernest Lever	//S//	9/28/2017
	Title: R&D Director	
REVIEWED BY: Tony Kosari	//S/	9/28/2017
	Title: Senior Engineer	
APPROVED BY: Ernest Lever	//S//	9/28/2017
	Title:	

Legal Notice

This information was prepared by Gas Technology Institute ("GTI") for the U.S. Department of Transportation Pipeline and Hazardous Materials Safety Administration.

Neither GTI, the members of GTI, the Sponsor(s), nor any person acting on behalf of any of them:

a. Makes any warranty or representation, express or implied with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately-owned rights. Inasmuch as this project is experimental in nature, the technical information, results, or conclusions cannot be predicted. Conclusions and analysis of results by GTI represent GTI's opinion based on inferences from measurements and empirical relationships, which inferences and assumptions are not infallible, and with respect to which competent specialists may differ.

b. Assumes any liability with respect to the use of, or for any and all damages resulting from the use of, any information, apparatus, method, or process disclosed in this report; any other use of, or reliance on, this report by any third party is at the third party's sole risk.

c. The results within this report relate only to the items tested.

Table of Contents

Page

Signature Page i

Legal Notice ii

Table of Contents iii

Table of Figures ix

List of Tables xix

Acknowledgements xx

Executive Summary 1

Impact from the Research Results 2

High Level Detail of Project Components 3

Report Structure 5

Technical Achievements of Project 10

Tools for Evaluating Risk 10

Non-Destructive Evaluation in Confined Spaces, Fitness for Service, Replacement Prioritization and Data integration 11

Summary 11

1. Rate Process Models (RPM) for Aldyl-A Piping Systems 12

Classic RPM, ASTM and ISO 13

Discussion of Coleman and Bragaw Papers 15

Reference RPM Data for Aldyl A 17

Material Ranking 21

Failure Stress Ranking 23

Failure Time Ranking 24

Operational Relative Ranking 24

The Impact of Pipe Inner Wall Surface Condition on Ranking Scores 28

Surface Condition Ranking Model V2.0 33

Applying the Surface Correlation Model 44

Surface Correlation Conclusions 45

Bi-directional Shift Factors 46

The molecular Structure of Polyethylene 46

Section 1 Bibliography 58

2. Stress Intensification Factors (SIF) for Aldyl-A Piping Systems 59

Brief Literature Review for Current Context 59

Failure Modes of Polyethylene 59

Variance of Stress in Pipes 60
Stress Intensification Factors (SIF) 63
SIF for Polyethylene Piping Systems 72
Combining SIF Distributions for Polyethylene Piping systems 78
SIF Summary 79
Section 2 Bibliography 80
3. Adjusted RPM Models for Use with Aldyl-A SIF 81
Adjusting the RPM Model for Internal SIF 81
Summary 85
4. Bayesian Network Model for Aldyl-A Lifetime Expectancy 86
Introduction 86
Bayesian Network Theory 86
Pipeline Risk Modeling 87
Graphical Analysis for Risk Factors 87
Bayesian Network Quantification for Aldyl A Risk Assessment 92
Prior Stress Intensification Factor (PSIF) 93
Impingement Stress Intensification Factor (ImpSIF) 100
Operational Stress Intensification Factor (OSIF) 106
Risk Score 110
Sub-networks Assembly 111
Bayesian Network Testing and Validation 112
Risk Score Validation 112
Rate Process Model Validation 113
Remaining Useful Life Prediction 114
Conclusions and discussions 116
Section 4 Bibliography 117
5. Structured Light Method for Internal Inspection of Aldyl-A Pipe 118
Introduction 118
Principle of 3D optical inspection system 118
Passive Acquisition 120
Multiple viewpoints 121
Single view points 121
Active Acquisition 122
Active-stereo 122
Laser Scanning 123
Structured Light 124
Multi colors multi rings approach: 124
Colored single shot patterns: 125
Projector pattern 125

Reconstruction algorithm 125
Image enhancement 125
Segmentation: 126
Color sequence detection: 126
Experimental 127
Defects and cracks in the pipes 127
Scanning without light source 129
Scanning with small view angle camera 131
3D image reconstruction 133
Scanning with large view angle camera 135
Scanner and the Sensor in Opposite Direction 138
Deformation Detection 139
Eccentricity Calculation 139
Eccentricity of a deformation free area 141
Eccentricity of the Area with Squeezed Parts 141
Misalignment Correction 142
Raspberry Pi based scanner: 144
Motion detection and distance estimation: 145
Multi-Rings Multi-Colors Scanning 148
Simulation environment: 148
Defects database: 153
Projector prototype III: 155
Miniaturized sensor (Prototype IV): 158
Structured light module: 160
Camera: 162
Triangulation in the scanner 163
Camera Calibration 163
Experimental results 164
Damage classification and data reduction 168
Classification results based on HOG descriptor 171
(1) Simulation pipe data classification 171
(2) Experimental pipe data classification 172
Pipe data convolutional neural network (CNN) classification model 173
Background 174
Pipe data preprocessing procedure 174
Pipe data classification results 175
Conclusions 179
Future work 180
Section 5 Bibliography 181
6. Damage Detection from Optical Imaging, Creep-Crack Growth Prediction, and Condition-based Maintenance Framework for Aldyl-A Pipes 183
Introduction 184
Imaging reconstruction and damage classification 186
Naïve Bayes classifier 187

Pipe imaging data simulation and classification 188
Pipe simulation 188
Feature calculation and classification 189
Comparison with neural network (NN) 190
Neural Network with extracted feature data 190
2D feature image as input 191
Classification using real image 193
Image reconstruction and damage classification conclusion 195
Bayesian/maximum entropy network 196
Introduction to maximum entropy 196
Maximum entropy in Bayesian network classifier 197
Demonstration example for Bayesian/maximum entropy network 200
Bayesian/maximum entropy network conclusion 202
Creep crack growth prediction 202
Model development 202
Calibration and validation 204
Uncertainty and reliability prediction 207
Approximated creep crack growth prediction conclusion 209
Maintenance framework 209
Maintenance framework formulation 210
Generating probability transition matrix 211
Illustrative example 212
Maintenance decision optimization 213
Consequence cost 214
Dynamic maintenance framework 216
Maintenance framework conclusion 220
Conclusions and future work 221
Future work 222
Section 6 Bibliography 224
7. Enterprise Decision Support System (EDSS) Framework 227
Introduction 228
Probabilistic Decision Support System Design 229
Semantic Ontology 229
Probabilistic Enterprise Decision Support System 231
Data Entry 236
Simulation 236
Insights Interface 237

Query 237
Data Communication 238
Probabilistic Decision Support System Development 244
Smart Form 244
Methodology 245
Root Cause Analysis Process 245
Fault Tree Analysis 245
Decision Tree 246
Gathering Context Sensitive Information for Risk Models 250
Eliminating the “Other” Category 250
Data Driven Risk and Consequence Models 251
Involving Standards Organizations 251
Benefits of Proposed Smart Form 251
Development 253
Components 256
User Personas 256
a. Field Technician 256
b. Manager/Business Analyst 257
c. Configuration Developer 257
Specifications 257
Functionality 257
User Interfaces 258
Performance 259
Operations & Maintenance 260
Systems Integration 260
Security 260
Data Management 261
Use Cases 261
Network Risk Assessment 261
Cognitive Field Data Collection 262
Field Collection Activities Use Case 263
Workflow Management 264
Design 265
Configuration Management 265
Situational Reasoning 266
Component Architecture 267
Model Management Techniques 268
Activity Logging and Provenance 268
Data Management 269
Aldyl-A Pipeline Network Model 270
User Interface and Integration with ESRI/ArcGIS 271
Geospatial View 271
Gains Chart 271
SmartForm Ontologies 272
Standard and 3rd Party Ontologies 272
Keyhole Inspection Form, First Response Form, Audit Report Form 273
Activity Diagrams 276
Activity Diagram – Create a tracking record for keyhole data gathering inspections 276

Activity Diagram – Create a tracking record for first response data gathering 277
Activity Diagram – Create a tracking record for audit event data gathering 278
Deployment and Testing 280
Test Data Preparation 280
Test Case Preparation.....281
Testing of Keyhole Inspection Form 282
Functional Testing 282
System Integration Testing 286
Validation Testing 286
Risk Calculation Testing 289
Testing of First Response Form 291
Functional Testing 292
System Integration Testing 292
Validation Testing 292
Testing of Audit Report Form.....298
Functional Testing 298
System Integration Testing 299
Validation Testing 299
Query Testing.....305
Conclusions 305
8. Feasibility of Using a Short-Term Pressure Test to Determine Fitness for Service in the Context of Slow Crack Growth 306
Understanding Highly Localized Ductile Creep Failure (SCG).....309
SCG Summary 336
Ductile Tearing Summary 336
Pressure Test Guidelines 337
Short-term Pressure Test Conclusion.....347
Section 8 Bibliography 348
9. Recommendations 349
10. Future Work 352
Fitness for Service 352
Decision Support systems.....353
Sensing Damage via Internal Inspection 354
List of Acronyms 356
References 358

Table of Figures

Page
Figure 1-1. Sequence of topics for discussing the RPM in the context of FFS.....12
Figure 1-2. Linear relationship of $\log_{10}$(time to failure) as a function of stress [5].....16
Figure 1-3. DuPont Control Aldyl-A at 23°C (73.4°F) Reference Temperature 20
Figure 1-4. Data Set 5 Points at 15°C Showing Correspondence with DuPont LDIW Data 20
Figure 1-5. The General RPM Model as Described by ISO 9080 21
Figure 1-6. Relative Ranking Bands for Material RPM Performance Relative to DuPont Reference Data.....23
Figure 1-7. Relative Ranking Bands for Failure Stress Relative to Absolute Hoop Stress in psi 24
Figure 1-8. Relative Ranking Bands for Failure Time Relative to Absolute Time to Failure in years.....25
Figure 1-9. Material RPM Relative Ranking for Data Set 1 @ 15°C 26
Figure 1-10. Failure Stress Relative Ranking for Data Set 1 @ 15°C.....26
Figure 1-11. Failure Time Relative Ranking for Data Set 1 @ 15°C 27
Figure 1-12. Operational Relative Ranking for Data Set 1 @ 15°C 27
Figure 1-13. Initial Surface Feature Model for Predicting “Risk” Scores 28
Figure 1-14. SEM image showing Crystals, Rods and Micro-crack 30
Figure 1-15. SEM Image Showing Dimples.....30
Figure 1-16. Revised Surface Feature Model.....33
Figure 1-17. Performance of Surface Feature and Thermal Properties Model for Material Ranking Category Prediction.....35
Figure 1-18. Probabilistic prediction of RPM performance band for data point 1 36
Figure 1-19. Probabilistic prediction of RPM performance band for data point 2 36
Figure 1-20. Probabilistic prediction of RPM performance band for data point 3 37
Figure 1-21 Probabilistic prediction of RPM performance band for data point 4 37
Figure 1-22. Probabilistic prediction of RPM performance band for data point 5 38
Figure 1-23. Probabilistic prediction of RPM performance band for data point 6 38
Figure 1-24. Probabilistic prediction of RPM performance band for data point 7 39
Figure 1-25. Probabilistic prediction of RPM performance band for data point 8 39
Figure 1-26. Probabilistic prediction of RPM performance band for data point 9 40
Figure 1-27. Probabilistic prediction of RPM performance band for data point 10 40
Figure 1-28. Probabilistic prediction of RPM performance band for data point 11 41
Figure 1-29. Probabilistic prediction of RPM performance band for data point 12 41
Figure 1-30. Probabilistic prediction of RPM performance band for data point 13 42

Figure 1-31. Probabilistic prediction of RPM performance band for data point 14 42
Figure 1-32. Probabilistic prediction of RPM performance band for data point 15 43
Figure 1-33. Probabilistic prediction of RPM performance band for data point 16 43
Figure 1-34. Highlighting Correctly-Predicted Reference Categories for Three Example Points 45
Figure 1-35. Polyethylene Structure and Potential Energies for Rotation of C-C Bond [12] 46
Figure 1-36. Unit Cell Structure of PE Crystallites [12] 47
Figure 1-37. Solidification model of Crystallization Process [13].....47
Figure 1-38. Left-Relaxation Mechanisms in PE, Upper Right – Mechanical α-Process Loss Tangent as a Function of Frequency, Lower Right - Mechanical α-Process Storage Modulus as a Function of Temperature [12] 48
Figure 1-39. Site-model as described by Bower [13] pp 148 50
Figure 1-40. Comparison of DTMA and Popelar Horizontal Shift Factors in Form Presented in [11].....53
Figure 1-41. Horizontal and Vertical Activation Energies of Aldyl A (PE2306) Pipe Batches Measured by DTMA. Energy in cal/mol 54
Figure 1-42.Horizontal Activation Energies for Modern and Aldyl A Materials Relative to Equivalent Popelar Activation Energy. Energy in cal/mol 55
Figure 1-43Vertical Activation Energies for Modern and Aldyl A Materials Relative to Equivalent Popelar Activation Energy. Energy in cal/mol.....55
Figure 1-44. 80°C Validation Test time/pressure Combinations for Several Materials to Validate a 20-year Minimum Lifetime at 23°C and an Operating Pressure of 55 psig.....56
Figure 1-45. Plot of RPM analysis results.....57
Figure 2-1. Schematic illustration of the failure behavior of pressurized PE 60
Figure 2-2. Aldyl A Data Sets for Deriving Empirical SIF 64
Figure 2-3. PDF for Empirically Derived SIF for Ductile Failures 64
Figure 2-4. CDF for Empirically Derived SIF for Ductile Failures 65
Figure 2-5. FEM analysis of stress associated with die line grooves on internal diameter of MDPE pipe.....66
Figure 2-6. SIF measured in FEM as a function of internal pressure at 73.4°F.....66
Figure 2-7. Distribution of ductile failure SIF referenced to FEM results 67
Figure 2-8. Empirical SIF Distributions for Heavily Grooved Gas Distribution Pipe.....68
Figure 2-9. RPM Analysis for Empirical SIF Determination.....68
Figure 2-10. SIF Distributions for Various Aldyl A SCG Data Sets 69
Figure 2-11. Actual vs Predicted Plot for SIF and Reference RPM Model Lifetime Prediction ...70
Figure 2-12. Same as Figure 2-11, but Linear Axes 70
Figure 2-13. Actual to Predicted Failure Time Ratio 71
Figure 2-14. CDF for Actual to Predicted Failure Time Ratio.....71

Figure 2-15. PPDC April 2017 Data for Polyethylene 74
Figure 2-16. FEM models for saddle tee SIF evaluation 75
Figure 2-17. SIF for pipe as a function of BRF and P 76
Figure 2-18. Bending moment sign convention-positive concave up, negative concave down ..76
Figure 2-19. Relative effect of bending configurations on saddle tee assemblies 77
Figure 2-20. SIF for pipe, tee and coupling under bending 77
Figure 2-21. Wikipedia contributors. "Logistic distribution." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 15 Jul. 2017. Web. 78
Figure 3-1. Distribution of SIF in DuPont SCG Reference Dataset 82
Figure 3-2. True Stress vs Hoop Stress in DuPont SCG Reference Set 82
Figure 3-3. SIF Adjusted RPM model for DuPont Control SCG Dataset 83
Figure 3-4. Adjusted RPM Model with Prediction Limit Bands 84
Figure 3-5. Validation of Adjusted RPM Model and Standalone SIF Approach for 4 Independent Datasets with Different Intrinsic Defects and Loading Conditions 85
Figure 4-1. Example of Bayesian network with five nodes 86
Figure 4-2. Factors selected by SMEs 87
Figure 4-3. Initial Bayesian network graph selected by SMEs 88
Figure 4-4. Betweenness centrality of the Bayesian network graph 89
Figure 4-5. Degree centrality of the Bayesian network graph 89
Figure 4-6. Composite ranking of the Bayesian network graph 89
Figure 4-7. Composite ranking of the nodes in Bayesian network graph 90
Figure 4-8. Methods to derive the conditional probability tables in the Bayesian network 91
Figure 4-9. Preliminary conditional probability distributions 91
Figure 4-10. Ontology describing the models, subject matter expert's knowledge, historical observation, and Bayesian network 92
Figure 4-11. The modified overall Bayesian network 93
Figure 4-12. Data sets for SCG control, LDIW control, LDIW with indentation and LDIW with squeeze-off shifted to 73.4°F 96
Figure 4-13. The SIF distributions under different loading conditions 97
Figure 4-14. The schematics for computing the relative SIF (SIFR) 98
Figure 4-15. The procedures for computing the RUL and total failure time 98
Figure 4-16. Shift the LDIW data leftwards by $\Delta \boldsymbol{t}$ 99
Figure 4-17. The validation of predicted total failure time after different years in service 99
Figure 4-18. The sub-network for computing the prior SIF distribution 100
Figure 4-19. The geometry of indentation/impingement 101
Figure 4-20. The stress strain curve for Aldyl A 102

Figure 4-21. The Bayesian network for strain components.....103
Figure 4-22. The distributions of SIF from Bayesian network (a) and experiment (b) 104
Figure 4-23. The Bayesian network model for computing impingement SIF 105
Figure 4-24. Bending moment sign convention-positive concave up, negative concave down 107
Figure 4-25. The relative severity of bending configurations at 45 psig internal pressure, SDR 11 pipe.....108
Figure 4-26. The schematic representation of pipeline bending undervertical soil movement (a) and horizontal soil movement (b) 108
Figure 4-27. The resultant deformation of pipeline bending due to soil movement 109
Figure 4-28. The Bayesian network for node “Operational SIF”.....110
Figure 4-29. The Bayesian network for node “Risk Score” 111
Figure 4-30. The integrated Bayesian network for pipeline risk assessment 112
Figure 4-31. The reproduced prediction limits using Bayesian network 114
Figure 4-32. The final distributions for the log of the life expectancy (hrs) Mean life=53.1 years, (b) Mean life=21.7 years, (c) Mean life=8.7 years 115
Figure 5-1. Example of a structured light pattern projected onto an object [30] 119
Figure 5-2: Illustration of a 3D structured light measurement [31] 120
Figure 5-3. Classification of the optical sensors category.....120
Figure 5-4. Illustration of multiple viewpoints technique [31].....121
Figure 5-5. Illustration of generating 3D image by using shape from shading [28].....122
Figure 5-6. Illustration of active stereo with laser scanning technique [31] 123
Figure 5-7. Laser scanning technique using CCD camera and laser source [34].....124
Figure 5-8. structured light technique with a light projector and camera [35] 124
Figure 5-9. Projection patterns: a) rectangular, b) circular.....126
Figure 5-10: Matching process.....127
Figure 5-11. Pipes for the experimental work 127
Figure 5-12: The holes that have been introduced as defects 128
Figure 5-13. The linear cracks in the spacemen.....128
Figure 5-14. Pipe with a small diameter (1.85 inches) and damage section 129
Figure 5-15. Optical outer and inner view of the E41 damage of GTI pipe sample damaged area 129
Figure 5-16. Two different cameras installed inside the pipe 130
Figure 5-17. The different defect types that have been captured by using the camera.....130
Figure 5-18. The laser source used in the sensor prototype.....131
Figure 5-19. Simple camera and source bundle prototype 132
Figure 5-20. Some frames form the collected data: a) with no defects and b) with defect 132

Figure 5-21. Results after applying the thresholding 133
Figure 5-22. Reconstruction algorithm for single laser ring.....134
Figure 5-23. 3D reconstruction of the internal surface of the pipe 135
Figure 5-24. Fish-eye camera 136
Figure 5-25. The camera and the laser source installed in temporary prototype.....136
Figure 5-26. Complete ring images, a) full ring. b) ring with some deformation 137
Figure 5-27. The 3D reconstructed image with complete FOV 137
Figure 5-28. The prototype that has been used in this section 138
Figure 5-29. Illustration of one frame that has defects.....138
Figure 5-30. 3D reconstruction.....139
Figure 5-31: Illustration of the circle which has fixed radius.....140
Figure 5-32. Illustration of the ellipse [43].....140
Figure 5-33. Illustration of 20 frames (4.4 cm) from collected data 141
Figure 5-34. a) Illustration of one frame with a defect and there are two lines to emphasize the ring center, b) Final result of the eccentricity has been calculated for 71 frames (17.6 cm), c) Illustration of the semi-major radius of all 71 frames (17.6 cm).....142
Figure 5-35. The region without defects but the shape is not circle 143
Figure 5-36. a) The eccentricity of multi frames without misalignment correction, b) 3D reconstructed image of number of frames without applying misalignment correction 143
Figure 5-37. a) illustration of one frame that effected by misalignment and the shape is not a circle. b) The same frame after applying misalignment correction and the shape is closer to circle. c) 3D reconstructed image after applying misalignment correction, d) The eccentricity after applying misalignment correction 144
Figure 5-38. a and b Represents the Reconstruction Results on Raspberry Pi 146
Figure 5-39. Equal Sized Slices 28cm long a) 0-7 cm b) 7-14cm c) 14-21cm.....146
Figure 5-40. a) Object under test for distance calculation, b) 3D Rendered Scan with Crack Detection.....147
Figure 5-41. a) Experimental flow for distance calibration, b) Comparison between the actual distance and measured distance.....148
Figure 5-42. Structured light with a projector used to generate the [45].....148
Figure 5-43. a) Illumination pattern, b) Geometry after illumination, c) Segmented image.....149
Figure 5-44. 3D reconstruction results 150
Figure 5-45. Circular geometry of the simulation.....150
Figure 5-46. Simulation of a normal defect free pipe 151
Figure 5-47. Simulation with a bend along the pipe.....151
Figure 5-48. Simulation of deformation along the pipe wall 152
Figure 5-49. Simulation of bump in the pipe wall.....152

Figure 5-50. Simulation of a hole in the pipe wall 153
Figure 5-51. The lenses stacked inside the prototype 155
Figure 5-52. a) The slide for the projector, b) Projected pattern 155
Figure 5-53. Projector and the camera connected to each other 156
Figure 5-54. a) the projection on the smooth surface. b) the prototype with blocked edges.....157
Figure 5-55. a) Illustration of one frame without defects. b) Illustration of a frame with squeezed part 157
Figure 5-56. Illumination of the white color has been reduced.....157
Figure 5-57. 3D reconstructed image with and without the defects.....158
Figure 5-58. Reconstructed frames that show the squeezed parts inside the pipe 158
Figure 5-59. proposed schematics for sensor design 159
Figure 5-60. 3D schematic of the light module 161
Figure 5-61. Ray tracing of the lenses.....161
Figure 5-62. a) Condenser lens, b) Objective lens 162
Figure 5-63. Assembly of camera and light module.....162
Figure 5-64. Triangulation of the system inside the pipe 163
Figure 5-65. Lens distortion removal, a) Before, b) After 164
Figure 5-66. One side of the light module cover 164
Figure 5-67.Light module internal components.....165
Figure 5-68. a) Projector final assembly, b) Projection of multiple rings 165
Figure 5-69. Scanner final assembly 166
Figure 5-70. a) Prototype IV scanning system, b) a single frame from the system 166
Figure 5-71. Reconstruction of a single frame 167
Figure 5-72. 3D reconstruction of pipe surface from multiple frames.....168
Figure 5-73: The final version of the scanning tool and a separate image of the light module .168
Figure 5-74. HOG descriptor implementation scheme.....170
Figure 5-75. Four types of damages 171
Figure 5-76. Four normalized histograms in one block of one simulation pipe image 172
Figure 5-77. The classification results of the test simulation samples.....172
Figure 5-78. four normalized histograms in one block of one experimental pipe image 173
Figure 5-79. classification results of the experimental samples 173
Figure 5-80. image rotation 175
Figure 5-81: CNN structure 175
Figure 5-82. segment result of one image 177
Figure 5-83. classification results of a consecutive sequence of frames.....178

Figure 6-1. Graphic model for a Naïve Bayes classifier with one class node and four feature nodes 187

Figure 6-2. Four types of common damage in pipe: dent (top left), slit (top right), squeeze-off (bottom left) and impingement (bottom right).....188

Figure 6-3. . Sample of the four types of damage: dent (top left), slit (top right), squeeze-off (bottom left) and impingement (bottom right).....189

Figure 6-4. Average accuracy vs. training data size with simulated feature data 190

Figure 6-5. Average accuracy for Neural Network using simulated data.....191

Figure 6-6. Image processing steps for extracting feature image. (a) the original frame for an impingement, (b) the dilation helps form a closed contour, (c) the closed circle is filled, (d) extracting the filled graphs from a round, (e) rotating the isolated damage into the same direction and (f) pooling to enlarge the damage and eliminate useless data.....192

Figure 6-7. The accuracy vs. training data size using a refined Gaussian node 193

Figure 6-8. The prototype camera.....194

Figure 6-9. Damaged pipes with slit (left), dent (middle) and impingement (right) 194

Figure 6-10. Two good reconstruction results 195

Figure 6-11. The average classification accuracy vs. training size using real data 195

Figure 6-12. Average accuracy for NB and BEN 201

Figure 6-13. The comparison of the updated probability distribution with and without constraint 202

Figure 6-14. Schematic plot of the crack at a notch root and the semi-circular geometry of a crack 203

Figure 6-15. Experimental data shifted to 23°C (73.4°F) 205

Figure 6-16. The geometry of the pipe 205

Figure 6-17. The SEM image showing an initial crack 206

Figure 6-18. Stress-Life curve for SCG data and its confidence bond against the lab data 206

Figure 6-19. The prediction of stress-life curve for (a) LDIW, (b) LDIW indent and (c) Squeeze-off.....207

Figure 6-20. The crack growth curve for SCG and its confidence bond at 1000 psi.....208

Figure 6-21. The empirical CDF and (b) PDF of the crack length distribution 208

Figure 6-22. The empirical CDF of the error comparing with the CDF of the MC sample error for a) SCG, b) LDIW, c) LDIW indent and d) Squeeze-off.....209

Figure 6-23. Illustration of the crack growth evolvement.....212

Figure 6-24. The visualization of maintenance plan for (a) repair and (b) replacement.....214

Figure 6-25. The cost and failure probability prediction for each group 216

Figure 6-26. The original and updated probability of Kt 217

Figure 6-27. The maintenance plan with updating at the end of the 4th time step (a) and (b), comparing with the plan without updating (c) and (d) 218

Figure 6-28. Comparison of cost for maintenance for each period 218
Figure 6-29. Continuous updating for parameter Kt.....219
Figure 6-30. The effect of continuous updates on periodical maintenance cost.....220
Figure 6-31. Total cost for maintenance in 10 period vs the number of updates.....220
Figure 7-1. Example of ontology representing threats leading to pipeline failure 229
Figure 7-2. Illustration of ontology describing causality in gas distribution pipeline.....230
Figure 7-3. Ontology describing the models, subject matter expert's knowledge, historical observation, and Bayesian network.....230
Figure 7-4. Bayesian foundational ontology 231
Figure 7-5. Conventional ontology.....231
Figure 7-6. Conceptual design of Enterprise Decision Support System.....232
Figure 7-7. Regulations and Standards guiding decision making process in EDSS.....232
Figure 7-8. Use case diagram of EDSS 233
Figure 7-9. 3rd party integration of EDSS platform in their services.....234
Figure 7-10. Architecture of Enterprise Decision Support System 235
Figure 7-11. Simulation model toolboxes with access to heterogeneous data sources.....237
Figure 7-12. Mockup version of insights interface dashboard.....237
Figure 7-13. Supported Query Types 238
Figure 7-14. Data flow between user interface, database, and model modules in EDSS.....239
Figure 7-15. Basic Workflow Architecture of Smart Forms 244
Figure 7-16. Example of a fault tree 245
Figure 7-17. Full FTA for leaks associated with mechanical fittings.....246
Figure 7-18. Fault Tree depiction of Decision Tree.....247
Figure 7-19. Decision Tree depiction of the FTA 247
Figure 7-20. Entry level for the "Leak Associated with Mechanical Fitting" decision tree 249
Figure 7-21. Moving along the decision tree towards the root cause 250
Figure 7-22. Mind Mapping Diagram of Smart Form Application 252
Figure 7-23. High level systems architecture developed for the Smart Forms Platform 254
Figure 7-24. Configuration of forms to specific workflows.....255
Figure 7-25. Example sequence of states in a SmartForms application, effectively configured in the Configuration Manager.....266
Figure 7-26. Example sequence of states in a SmartForms application, effectively configured in the Configuration Manager.....268
Figure 7-27. Layer population techniques 270
Figure 7-28. Example of a Gains Chart 272
Figure 7-29. Test gas distribution network.....281

Figure 7-30. Screen captures of Keyhole Inspection Form mobile application.....285
Figure 7-31. Annual Report Form login page and main section.....299
Figure 8-1. Kiefner [89] Figure 1. 308
Figure 8-2. Design of Universal Test Vessell (UTV), with and without SIF notch.....309
Figure 8-3. Ductile failure in un-notched UTV.....310
Figure 8-4. Ductile failure initiated at the stress riser notch 311
Figure 8-5. Early detection of wall breach due to SCG 312
Figure 8-6. Constitutive model for FEM analysis of unimodal MDPE.....313
Figure 8-7. Engineering and True Stress / True Strain curves from the constitutive model.....314
Figure 8-8. Draw ratio and onset of strain hardening.....315
Figure 8-9. FEM analysis of the stress regime at the UTV notch tip 316
Figure 8-10. Zooming in to better visualize the notch tip stress field.....317
Figure 8-11. Computerized X-Ray Tomography image of damage propagation from the notch tip in a first generation unimodal MDPE UTV 317
Figure 8-12. Computerized X-Ray Tomography section through UTV showing multiple damage fronts propagating radially outwards from the notch tip 318
Figure 8-13. Specimen outer wall showing breach 319
Figure 8-14. Inner surface showing stress cracking 319
Figure 8-15. Fracture surface showing inner wall at top and outer wall at bottom.....320
Figure 8-16. Detail of fracture surface – inner region.....320
Figure 8-17. Detail of fracture surface – inner third region.....321
Figure 8-18. Detail of fracture surface – outer third region 321
Figure 8-19. Detail of fracture surface – wall breach region.....322
Figure 8-20. Increasing damage tip stress as a function of test pressure and depth of damage in an externally notched pipe 323
Figure 8-21. Decreasing triaxiality (loss of constraint) at the damage tip as a function of test pressure and depth of damage in an externally notched pipe.....324
Figure 8-22. Dependence of critical strain at break and fracture mode of stress triaxiality per Bao [90] (PhD Thesis, MIT 2003).....325
Figure 8-23. Void and fibril formation in polymers under stress triaxiality. Kausch, Polymer Fracture, Fig 9.17 a, p 340 [95].....325
Figure 8-24. Stress triaxiality contours around UTV notch. Outer closed contour is for triaxiality = 1/3, and demarks the region where hoop stress dominates (outside the closed contour). 326
Figure 8-25. Stress triaxiality contours around UTV notch. Outer closed contour is for triaxiality = 1, and demarks the region where void formation is possible (inside the closed contour). 326
Figure 8-26. Stress triaxiality contours around UTV notch. innermost closed contour is for triaxiality = 3, and demarks the region where void formation is highly likely to occur (inside the closed contour). 327

Figure 8-27. Same as Figure 8-26 – accentuating the stress levels at the maximum triaxiality region.....327
Figure 8-28. External Notch, 40% depth, 110 psig internal pressure (550 psi hoop stress).....328
Figure 8-29. External Notch, 70% depth, 110 psig internal pressure (550 psi hoop stress).....328
Figure 8-30. External Notch, 70% depth, 50 psig internal pressure (250 psi hoop stress).....329
Figure 8-31. Variations in critical strain at break in multiple unimodal MDPE specimens, cut from the same pipe, under the same nominal tensile test conditions.....330
Figure 8-32. Variation in the energy density for the tensile tests shown in Figure 8-31.....330
Figure 8-33. MDPE - comparison of creep rates under load controlled long-term tensile creep @ 90°C, 5 MPa.....331
Figure 8-34. Left specimen blue curve above, Right specimen red curve above 332
Figure 8-35. Creep Surface for MDPE 334
Figure 8-36. Creep Response of MDPE Pipe at 4313 h of Load Controlled Creep @ 90°C. ...335
Figure 8-37. Energy Density and Strain at Break for MDPE Pipe 336
Figure 8-38. Process for establishing pressure test guidelines.....337
Figure 8-39. Reference data sets shifted to 73.4°F per steps 1,2 and 3 of the process for determining pressure test guidelines.....338
Figure 8-40. FEM analysis of stress associated with die line grooves on internal diameter of MDPE pipe.....339
Figure 8-41. SIF measured in FEM as a function of internal pressure at 73.4°F.....339
Figure 8-42. Distribution of SIF calculated in steps 4,5,6 and 7 of the process.....340
Figure 8-43. Distribution of SIF in SCG data set.....340
Figure 8-44. Actual failure times vs predicted failure times, red points non-conservative 341
Figure 8-45. Correlation of model results to actual results mean=0.9.....341
Figure 8-46. Distribution of propagation time to fail time ratio.....343
Figure 8-47. Distribution of calculated failure times to actual failure times ratio.....344
Figure 8-48. Cumulative Density Plot for ratio of calculated failure times to actual failure time345
Figure 8-49. Over-pressure test @ 68°F and 150 psig internal pressure.....346
Figure 8-50. Over-pressure test @ 68°F and 400 psig internal pressure.....346

List of Tables

Page

Table 1-1. DuPont Control Model Parameters for 3 Parameter ISO 9080 Model 21

Table 1-2. Data for Initial Surface Feature Model 29

Table 1-3. Data for Expanded Surface Feature and Thermal Characteristics Model 31

Table 1-4 – Improved Surface Condition Ranking Model: 14 correct predictions, 2 conservative errors; 87.5% success rate 34

Table 1-5 – Data for Evaluating Use of the Model (Table sorted by stress) 44

Table 1-6. Activation Energies for α relaxation process in polyethylene gas distribution pipe as measured by DTMA 56

Table 2-1. Aldyl A RPM models for Region A and Region B 59

Table 2-2. Stand-alone SIF for Polyethylene Piping Systems 72

Table 2-3 Maupin and Mamoun summary of failure categories 73

Table 2-4. Correlation coefficients for pipe under bending with and without fittings 76

Table 2-5. SIF Distributions for Polyethylene Piping systems 79

Table 3-1. SIF Adjusted DuPont SCG Control Model Parameters for 3 Parameter ISO 9080 Model 83

Table 3-2. Adjusted SIF for use with Adjusted RPM Model 84

Table 4-1. The summary for node types and states 93

Table 4-2. Different combinations of loading conditions 100

Table 4-3. The type and states for each node 105

Table 4-4. Regression coefficients for each configuration 107

Table 4-5. The risk score comparison between logistic regression and Bayesian network 112

Table 4-6. The context conditions for 3 scenarios 114

Table 4-7. The context conditions for 10 scenarios 115

Table 4-8. The median and quantile life prediction for those 10 scenarios (years) 116

Table 5-1. Database defects types 153

Table 6-1. Random generator for demonstration data 200

Table 6-2. Uncertainty quantification of the model parameters for SCG group 207

Table 6-3. Nomenclature for the maintenance decision framework 210

Table 6-4. The condition stage 211

Table 7-1. Data for creation of gains charts 272

Table 7-2. Keyhole Inspection Form and First Response Form 273

Table 7-3. Technical Specification of the Deployment Environment 280

Table 7-4. Data before Keyhole inspection collects new information 290

Table 7-5. Latest data from Keyhole inspection.....290

Table 7-6. Mean life prediction of pipeline segments before and after data is collected with keyhole inspection form 290

Table 7-7. Risk after a number of years 291

Table 8-1. Creep Data for 2" SDR 11 MDPE Gas Distribution Pipe. Color Coded Per Pipe (3 pipes with multiple specimens from each) 332

Acknowledgements

This project was a large and ambitious project that required close coordination between Gas Technology institute (GTI), Arizona State University (ASU), University of Colorado Denver (UCD)/ Michigan State University (MSU), and other entities. There were many people in each of these organizations that played large and small parts in this effort. Their diligence and dedication to task was very much appreciated and we would be remiss not to mention them by name and offer our sincerest appreciation and thanks for their contributions.

GTI:
Tishun Peng, Saurav Acharya, Dennis Jarnecke, Daniel Ersoy, Shannon Katcher, Tom Purk, Khalid Farrag, Oren Lever, Pete Mulligan, Tony Kosari, Travis Cottle, Joe Irwin
ASU:
Yongming Liu (faculty), Yuhao Wang (PhD student), Sonam Dahire (PhD student), Songyan Shi (undergraduate)
UCD, MSU:
Yiming Deng (faculty), Mohand Alzuhiri (PhD student), Peipei Zhu (PhD student), Samea Elfaid (MS student)
Technical Advisory Panel:
Ed Newton, Andy Benedict, Bryan Hauger, Jim Merritt, Gene Palermo, Aaron Forster
SmartCloud Inc.:
Mike Barnett, Richard Anderson, Kim Mayyasi

Executive Summary

The primary objective of this project was to provide an integrated set of quantitative tools that provide a structured approach to evaluating the latent risk in vintage polyethylene pipes, such as Aldyl-A, that are common in gas distribution systems. Secondary objectives were: first, to provide a fitness for service approach that can support replacement prioritization; second, utilize data from multiple sources such as in ditch condition assessment and leak records; third, to provide a means to access the pipe in a congested urban environment.

The primary objective was realized through the construction of a hybrid causal Bayesian network that allows the synthesis of subject matter expertise with well understood physical causation pathways to failure. These logical structures inherently integrate disparate data sources into a coherent framework for decision support that is readily updated as new information becomes available.

From the outset, the project was structured around developing software agents that are modular and can be integrated into any decision support framework. The framework chosen in this project was a commercially available Artificial Intelligence (AI) platform that has built in ability to deal with any given business logic. The Bayesian networks developed in this project were successfully integrated into the AI framework and used to assess risk dynamically on a synthetic geographical dataset with over 7,000 data elements.

The secondary objectives were realized through several project components that are described next:

- Intelligent data entry forms were developed and integrated into the AI framework to capture disparate data and route the data to the relevant models. Bayesian updating of the model output, after this data was input via the forms, was demonstrated.
- Several variants of prototype endoscopic structured-light cameras were developed to measure pipe geometry from inside the pipe. The cameras developed are small enough to insert into 1" IPS diameter tubing and are suitable for integration into existing keyhole methods for accessing pipelines in congested urban environments. The project demonstrated that these cameras can identify several common geometric features consistent with pipeline anomalies known to introduce significant stress risers that drive pipeline failure via Slow Crack Growth (SCG). Feature recognition and data reduction methods were developed for the massive amounts of data acquired by the endoscopic tool
- A Stress Intensification Factor (SIF) based approach was developed for fitness for purpose determination. Two approaches were developed: reliability based maintenance models based on a fracture mechanics damage propagation approach were used to simulate the cost effectiveness of various repair/replace approaches, and a Bayesian network approach that integrates pipeline configuration, loading

conditions and material models was used to predict expected lifetime for pipeline segments. Both approaches were validated against historic data sets and shown to be equally effective. The material models were used to show that there is no effective pressure test approach to support pipeline replacement prioritization for polyethylene pipe. The lifetime expectancy of components given the available data inputs into the detailed material models developed in this project is the only viable option for enhancing decisions aimed at ensuring system integrity.

Impact from the Research Results

The insights developed from this body of work were not available to regulators and operators prior to this work, which has provided a structured set of tools for assessing the fitness for service of vintage polyethylene pipelines in gas distribution systems. The models developed in this project are comprehensive probabilistic risk models that can be fully integrated into enterprise decision support systems and used to prioritize replacement programs, provide system integrity reports and assist operators in identifying future integrity related problems in their systems. The prototype endoscopic tools have the potential to fundamentally change how integrity data can be gathered to feed improved risk models for vintage pipeline systems. A detailed list of potential follow-on work is provided in the body of the report. The results of this project will be publicly disseminated through several papers in industry journals and conferences. The scope of the project and interim results were presented at Plastic Pipes XVIII, Berlin, Germany, September 2016 and published in the conference proceedings. A follow-on paper with results will be submitted to Plastic Pipes XIX, Las Vegas, USA, September 2018.

Additional publications related to this body of work include:

1. Yuhao Wang. Yongming Liu, "Probabilistic life prediction and prognostics-based maintenance optimization for gas pipelines". annual conference of the prognostics and health management society 2017, St. Petersburg, Florida, October, 2017. (accepted)
2. Yuhao Wang, Yongming Liu, Tishun Peng, Ernest Lever. "Probabilistic life prediction of plastic pipes using an equivalent crack growth model". 2018 AIAA SciTech conference – nondeterministic approach, Kissimmee, FL, January, 2018. (accepted)
3. Yuhao Wang, Yongming Liu. "A Novel Bayesian Entropy Network for Probabilistic Damage Detection and Classification". 2018 AIAA SciTech conference – nondeterministic approach, Kissimmee, FL, January, 2018. (accepted)
4. Yongming Liu, Yiming Deng "Fast automatic anomaly characterization and risk management in gas pipelines", DOT PHMSA 2016 Research Forum – Invited talk, Cleveland, OH, October, 2016.
5. Yuhao Wang, Yongming Liu. "Reliability-based maintenance optimization for aging gas pipeline system with Bayesian updating". Reliability Engineering and System Safety, 2017. (under preparation)

6. Mohand Alzuhiri, Yiming Deng, "Phase extraction algorithm based fast 3D reconstruction for structured light sensors", to be submitted to Research in Nondestructive Evaluation, 2018 (journal)
7. Peipei Zhu, Mohand Alzuhiri, Yiming Deng, " Data fusion for plastic pipeline damage detection using machine learning algorithms", to be submitted to NDT & E, International, 2017 (journal)
8. Mohand Alzuhiri, Yiming Deng, "Structured Light Based Endoscopic Scanner for Small Diameter Gas Pipelines", 22nd ENDE Workshop, Saclay, France 2017, abstract published in ENDE 2017 proceedings (Sept. 2017)
9. Two published MS theses and one MS thesis to be defended December 2017

High Level Detail of Project Components

1. Stress risers are the most significant driver of failure in plastic gas-distribution pipe
    a. The scope of his project included a comprehensive review of existing data sets
        i. Historic DuPont data
        ii. Reference Rate Process Method (RPM) models were developed for all available data sets
        iii. Comprehensive testing of pipe exhumed after 20 and 40 years in service
    b. Correlations between features observable under a microscope and lifetime expectancy have been developed
    c. Equivalent Stress Intensification Factor (SIF) distributions have been developed for:
        i. Surface morphology
        ii. Low Ductile Inner Wall (LDIW)
        iii. Known loading conditions such as:
            1. Impingement
            2. Bending
            3. Squeeze-off
            4. Soil movement
            5. Ageing
2. It is possible to detect many significant stress risers from inside the pipe by measuring the internal pipe geometry
    a. This is the motivation for the endoscopic structured light tool development undertaken by Michigan State university (MSU)
    b. The method is compatible with keyhole methods already employed
    c. The project was successful in demonstrating that useful information can be gathered from compact structured-light tools and reduced to a useable format
3. A lifetime prediction approach, based on equivalent SIF and the control reference-RPM model for Aldyl-A, that can be applied to any combination of material and loading condition was developed

    a. Bayesian Networks (BN) were developed to capture material and loading combinations and output a predicted lifetime
    b. The BN output was validated against the reference RPM models
4. An alternative lifetime prediction approach utilizing fracture mechanics and damage propagation principles was developed by Arizona State University (ASU)
    a. There are many similarities in the GTI and ASU approaches as far as they both rely on SIF distributions.
        i. GTI utilizes the RPM
        ii. ASU utilizes a damage propagation approach
    b. Both methods provide acceptable predictions when validated against the reference data sets
5. The lifetime prediction models allow the use of Markov Chain Monte Carlo (MCMC) methods to evaluate the impact of various mitigation strategies on the reliability and cost of maintenance of an Aldyl-A gas-distribution system
6. Intelligent data collection approaches were developed to facilitate the capture of pertinent information, from a risk assessment perspective, from disparate data sources
    a. Basic SmartForms were developed for:
        i. First response forms
        ii. Keyhole data collection forms
        iii. Audit forms
    b. The data captured by the SmartForms was connected to the relevant models and Bayesian updating of the predicted risk profile was performed based on the new information
    c. These intelligent data collection forms are necessary for automated compilation of data needed to electronically submit annual and semi-annual pipeline integrity management program performance reports as specified in Advisory Bulletin ADB-07-01

Report Structure

Report Section	Topics Covered and Deliverables
Section 1: Rate Process Models (RPM) for Aldyl-A Piping Systems	**Topics:** 1. Rate Process Method a. Ranking methods b. Impact of inner wall surface on RPM 2. Polyethylene structure and molecular motion 3. Bi-directional shift factors
	Deliverables: 1. Reference RPM model for Aldyl A 2. Model for correlating surface condition to RPM ranking 3. Validated bi-directional shift factors for Aldyl A and other polyethylene materials
Section 2: Stress Intensification Factors (SIF) for Aldyl-A Piping Systems	**Topics:** 1. Failure mechanism of polyethylene 2. Variance of stress in polyethylene pipes 3. What is a SIF
	Deliverables: 1. SIF for various polyethylene piping configurations 2. Probability distributions for SIF in polyethylene piping systems with a quantification of uncertainty for the mean 3. Method for combining multiple SIF in a coherent manner

Report Section	Topics Covered and Deliverables
Section 3: Adjusted RPM Models for Use with Aldyl-A SIF	**Topics:** 1. Need to adjust the control RPM model to account for distribution of SIF included in the dataset 2. Need to calibrate standard SIF distributions for use with the adjusted RPM model 3. Validation of the adjusted RPM model and calibrated SIF **Deliverables:** 1. Universal model for SCG failure mode of Aldyl A pipe 2. Calibrated SIF for known pipe conditions 3. Validation of model and SIF over four independent reference datasets
Section 4: Bayesian Network Model for Aldyl-A Lifetime Expectancy Systems	**Topics:** 1. Causal mechanisms for generating stress fields in Aldyl A pipe 2. Methodology for addressing interacting causal factors 3. Calibration of Bayesian network to reference datasets **Deliverables:** 1. Comprehensive expert system for Aldyl A piping systems that addresses all interactions 2. Calibrated and validated Bayesian network for predicting lifetime expectancy of Aldyl A piping segments

Report Section	Topics Covered and Deliverables
Section 5: Structured Light Method for Internal Inspection of Aldyl-A Pipe	**Topics:** 1. Development of endoscopic structured light tool for internal inspection of gas distribution pipe 2. Algorithms for detecting and categorizing internal pipeline defects and geometry 3. Motion detection and distance measurement 4. Data reduction techniques **Deliverables:** 1. Working prototype of endoscopic structured light tool for internal inspection of gas distribution pipe 2. Algorithms for damage detection and classification 3. Data reduction methods
Section 6: Damage Detection from Optical Imaging, Creep-Crack Growth Prediction, and Condition-based Maintenance Framework for Aldyl-A Pipes	**Topics:** 1. Image reconstruction and damage classification 2. Bayesian maximum entropy network 3. Creep crack growth prediction 4. Maintenance frameworks **Deliverables:** 1. **Algorithms for image reconstruction** 2. **Working Bayesian maximum entropy network** 3. **Validated creep crack growth model for lifetime prediction for Aldyl A pipe segments** 4. **Framework for optimizing maintenance of Aldyl A piping systems i.e. repair/replace balance**

Report Section	Topics Covered and Deliverables
Section 7: Enterprise Decision Support System (EDSS) Framework	**Topics:** 1. Semantics of decision support framework for Aldyl A piping systems 2. Probabilistic decision support framework 3. Data entry 4. Simulation 5. Insights 6. Query 7. Communication **Deliverables:** 1. Functioning enterprise decision support framework that incorporates the Bayesian network developed in section 4 2. Smart forms for data entry 3. Insights and analytics framework 4. Data reporting framework
Section 8: Feasibility of Using a Short-Term Pressure Test to Determine Fitness for Service in the Context of Slow Crack Growth Systems	**Topics:** 1. Historic usage of short-term pressure tests in determining fitness for service of gas transmission pipe 2. Applicability to plastic gas distribution pipe 3. Localized damage propagation in plastic pipe **Deliverables:** 1. Conclusion that short-term testing is not a viable fitness for service test for plastic gas distribution pipe 2. Explanation of reasons why the method is not viable

Report Section	Topics Covered and Deliverables
Section 9: Summary and Recommendations	**Topics:** Summarizes technical achievements and recommends utilizing the methods developed in the project to transition to probabilistic, reliability based risk assessment methods. Recommends establishing a Joint Industry Program (JIP) to develop guideline for this transition. Recommends that the JIP oversee the commercialization and deployment of the structures light scanning tool that can be integrated into existing keyhole technologies
Section 10: Recommendations for Future Work	**Topics:** Lists future work effort in three areas: 1. Fitness for Service 2. Decision Support Systems 3. Sensing Damage via Internal Inspection The future work efforts are designed to enhance the methods developed, address potential weaknesses in the approaches and further integrate disparate efforts that all contribute to system risk and integrity.

Technical Achievements of Project

Tools for Evaluating Risk

The primary objective of this project was to provide an integrated set of quantitative tools that provide a structured approach to evaluating the latent risk in vintage polyethylene pipes, such as Aldyl-A, that are common in gas distribution systems.

Polyethylene pipes undergo constant creep due to the nature of the material. The underlying molecular mechanisms that enable creep are called relaxation mechanisms and they are constantly in action. Basic molecular motions occur thousands, or millions of times per second. If there is an external driving for that loads the polymeric structure, stresses will be developed in the material. These stresses give directionality to the random molecular motions that result in creep, and ultimately lead to failure of the polyethylene structure.

A primary deliverable of this project is a set of tools that define:

1. A Rate Process Method (RPM) model that defines the rate at which the polyethylene will creep.
    a. This rate is strongly dependent on temperature
    b. The rate is also dependent on the stress in the polyethylene structure that in turn depends on:
        i. The geometry of the component
        ii. The external loads acting on the component
2. The Stress Intensification Factors (SIF) that provide a simple means of translating the nominal hoop stress of the pipe, which is very easy to calculate, to true stress. Well defined and simple to apply SIF are essential to a workable risk evaluation method that utilizes simple, well-known parameters such as: pipe size, ambient temperature, system operating pressure, component configuration and other measurable installation characteristics to arrive at a true component stress. The SIF developed in this project, together with a single master RPM model underpin the lifetime prediction methods presented in this project.
3. The RPM model and SIF can be used as-is to perform risk assessments given system parameters, or they can be integrated into a tool that is capable of integrating all threat interactions into a composite risk score. A deliverable of this project is a Bayesian network that accomplishes this objective. A fully defined, calibrated and validated Bayesian network is defined in this project.

Non-Destructive Evaluation in Confined Spaces, Fitness for Service, Replacement Prioritization and Data integration

Secondary objectives of the project were: first, to provide a fitness for service approach that can support replacement prioritization; second, utilize data from multiple sources such as in ditch condition assessment and leak records; third, to provide a means to access the pipe in a congested urban environment. These objectives were technically realized through:

1. A non-destructive tool that is capable of measuring pipeline configuration from inside the pipe was developed and prototyped in this project. This structured light, endoscopic measurement tool is a major breakthrough in assessing gas distribution pipes as it allows the operator to measure pipeline geometry over large lengths of the pipe without excavating the entire pipe. This measurement of pipe geometry from inside the pipe allows identification of several critical defects such as: impingement, squeeze-off, fittings, sudden displacements of the pipe, pipe deformations and other defects that cause stress intensification. This direct measurement of features will allow accurate SIF to be assigned to pipe segments. This will allow proper classification of segment with regard to the anticipated stress fields that when plugged in to the RPM model will provide probability of failure over time. This likelihood of failure is a key component in determining the segments Fitness for Service (FFS).
2. A set of reliability based tools were developed that underpin optimization methods for comparing repair/replace strategies over multi-year timeframes. These methods are based on robust damage propagation methods that were calibrated and validated against historic reference data. It was demonstrated that the Monte Carlo simulations that these tool support are capable of evaluated multiple scenarios and providing guidance as to the most effective risk management strategy over time
3. The tools developed in this project were integrated into a commercially available Artificial Intelligence (AI) platform that is capable of merging multiple disparate data sources, running the various risk assessment tools and providing insights driven be sophisticated data analytics. Intelligent forms that facilitate in-field data gathering and regulatory reporting requirement were also developed and demonstrated and tested as part of this project

Summary

All of the project deliverables were met and tested via the components described above. A comprehensive set of tools that can be practically applied in multiple approaches, from simpe point applications, to enterprise wide decision support systems has been provided.

1. Rate Process Models (RPM) for Aldyl-A Piping Systems

This section discusses the use of the Rate Process Method (RPM) in the context of determining the Fitness for Service (FFS) of Aldyl A gas distribution pipe. The relevance of the RPM in determining FFS of plastic piping is through the usefulness of the method in determining the lifetime expectancy of a pipe under a given load at a given temperature. The typical usage of the RPM correlates hoop stress to time to failure [1, 2]. This normal usage of the RPM is not sufficiently accurate for determining FFS, so we have developed a set of approaches for improving the accuracy of RPM predictions. These approaches will be described in detail in Sections **2** and **3**. The discussion in this section will follow the sequence laid out in **Figure 1-1**.

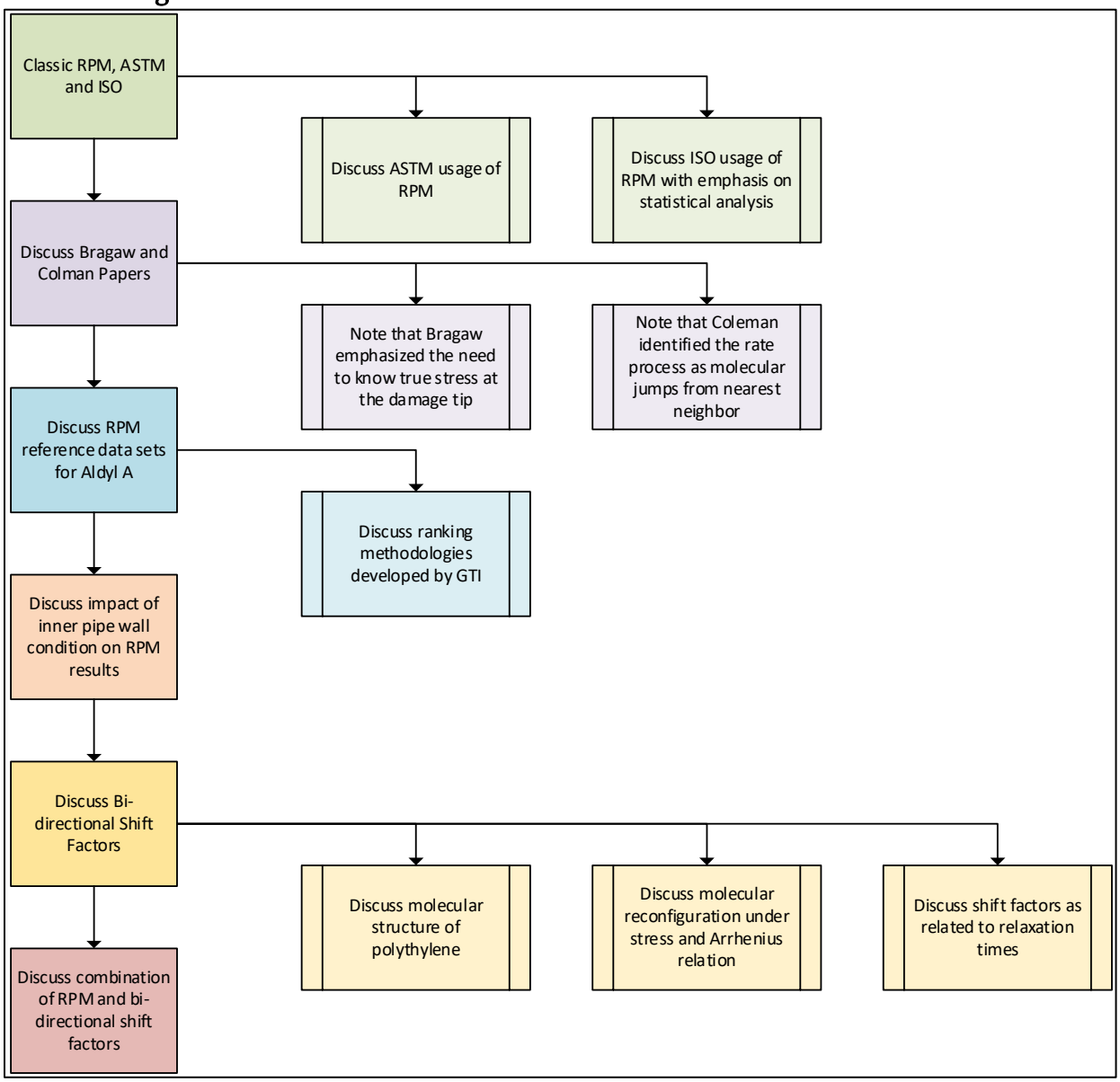


Figure 1-1. Sequence of topics for discussing the RPM in the context of FFS

Classic RPM, ASTM and ISO

Kurdziel and Palermo [1] provide a concise overview of the development of the RPM in the United Sates and how it became incorporated into ASTM standards. The insert below reproduces the relevant portion of the paper.

The Rate Process Method (RPM) [1]

The concept and mathematical basis for using the Rate Process Method for polyethylene (PE) pipe and fitting service projections was originally presented by Bragaw (2) (3). The Plastics Pipe Institute (PPI) Hydrostatic Stress Board (HSB) conducted an extensive evaluation of this and other methods for forecasting the effective long-term performance of PE piping materials. Basically, all these methods require elevated temperature sustained pressure testing of pipe where the type of failure is of the slit or brittle-like mode.

As a result of these studies, HSB determined that the three-coefficient Rate Process Method (RPM) equation provided the best correlation between calculated long-term performance projections and known field performance of several PE piping materials. It also had the best probability for extrapolation of data based on the statistical "lack of fit" test. Further validation of the Rate Process Method was made by comparing RPM projections for PE pipe and fittings obtained at elevated temperatures with actual room temperature laboratory failures for the same pipe and fittings (4).

Rate Process Method testing of pipe or fitting assemblies is conducted in accordance with ASTM D 1598, "Standard Test Method for Time-to-Failure of Plastic Pipe Under Constant Internal Pressure" (5). Fittings are joined to pipe using standard joining procedures. Conducting an RPM experiment requires a minimum of 18 to 20 specimens at various temperature/pressure conditions. As with any test protocol, increasing the number of specimens provides a higher confidence level in the failure mode validation and limits.

Using slit failure mode data points, one calculates the A, B and C coefficients for the following three-coefficient Rate Process Method extrapolation equation:

$$\text{Log}\, t = A + \frac{B}{T} + \frac{C\,\text{Log}\, S}{T}$$

Where:

t = slit mode failure time (hours)
T = absolute temperature (K)
S = stress (psi)

Once the A, B and C coefficients are determined, the RPM equation can be used for various performance projections (average failure time) at typical use temperature (average annual ground temperature) and stress conditions.

The RPM provides the means for not only validating the long-term performance capacity for corrugated HPDE pipe, but it provides a basis for assessing the manufacturer's quality assurance or quality control program. After establishing the RPM coefficients, an appropriate single-point elevated temperature stress rupture test may be established for quality purposes (6).

Mathematically, these RPM projections are sound. They are, however, not absolute and are subject to various experimental errors, unknown deviations and judgment factors. The calculations from the RPM equation are used in conjunction with other known mechanical, performance, and design factors specifically relating to corrugated HDPE pipe to validate the service life projections for these applications.

2. C. G. Bragaw, "Crack Stability Under Load and the Bending Resistance of MDPE Piping Systems", Seventh Plastic Fuel Gas Pipe Symposium, New Orleans, October 1980.
3. C. G. Bragaw, "Service Rating of Polyethylene Systems by the Rate Process Method", Eighth Plastic Fuel Gas Pipe Symposium, New Orleans, November 1983.
4. E. F. Palermo, "Rate Process Method as a Practical Approach to a Quality Control Method for Polyethylene Pipe", Eighth Plastic Fuel Gas Pipe Symposium, New Orleans, November 1983.
5. American Standards Testing and Materials (ASTM), ASTM D 1598, "Standard Test Method for Time-to-Failure of Plastic Pipe Under Constant Internal Pressure," 2004.
6. E. F. Palermo, "Using Laboratory Tests on PE Piping Systems to Solve Gas Distribution Engineering Problems", Tenth Plastic Fuel Gas Pipe Symposium, New Orleans, October 1987.

Kurdziel and Palermo point out that although the method is mathematically sound it is subject to various experimental errors, unknown deviations and judgement factors.
The ISO approach to the RPM is described in "ISO 9080:2012 Plastics piping and ducting systems -- Determination of the long-term hydrostatic strength of thermoplastics materials in pipe form by extrapolation" [2]. The sections covering the principles and use of the method are reproduced in the inserts below.

Principles [2]

The suitability for use of a plastics pressure pipe is first of all determined by the performance under stress of its material of construction, taking into account the envisaged service conditions (e.g. temperature). It is conventional to express this by means of the hydrostatic (hoop) stress which a plastics pipe made of the material under consideration is expected to be able to withstand for 50 years at an ambient temperature of 20 °C using water as the internal test medium. The outside environment can be water or air.

In certain cases, it is necessary to determine the value of the hydrostatic strength at either shorter lifetimes or higher temperatures, or on occasion both. The method given in this International Standard is designed to meet the need for both types of estimate. The result obtained will indicate the lower prediction limit (LPL), which is the lower confidence limit of the prediction of the value of the stress that can cause failure in the stated time at a stated temperature (the ultimate stress).

NOTE The MRS value (at 20 °C) is usually based on data obtained using water as the internal and external test medium. It is obvious that indeed all data are used for validation of regression curves at higher temperatures (e.g. 70 °C), including the data obtained with air as the external medium (e.g. at 110 °C).

This International Standard provides a definitive procedure incorporating an extrapolation using test data at different temperatures analysed by multiple linear regression analysis. The results permit the determination of material-specific design values in accordance with the procedures described in the relevant system standards. This multiple linear regression analysis is based on the rate processes most accurately described by log10(stress) versus log10(time) models.

In order to assess the predictive value of the model used, it has been considered necessary to make use of the estimated 97,5 % lower prediction limit (LPL). The 97,5 % lower prediction limit is equivalent to the lower confidence limit of the 95 % confidence interval of the predicted value. This convention is used in the mathematical calculations to be consistent with the literature. This aspect necessitates the use of statistical techniques.

The method can provide a systematic basis for the interpolation and extrapolation of stress rupture characteristics at operating conditions different from the conventional 50 years at 20 °C. Taking into account the extrapolation factors (see 5.1.4), the extrapolation time limit can go up to 100 years.

It is essential that the medium used for pressurizing the pipe does not have an adverse effect on the pipe. In general, water is considered to be such a medium.

Long consideration was given to deciding which variable should be taken as the independent variable to calculate the long-term hydrostatic strength. The choice was between time and stress.

The basic question the method has to answer can be formulated in two ways as follows.

a) What is the maximum stress (or pressure) that a given pipe system can withstand at a given temperature for a defined time?

b) How long will a pipe system last when subjected to a defined stress (or pressure) at a given temperature?

Both questions are relevant.

If the test data for the pipe under study does not show any scatter and if the pipe material can be described perfectly by the chosen empirical model, the regression with either time independence or stress independence will be identical. This is never the case because the circumstances of testing are never ideal nor will the material be 100 % homogeneous. The observations will therefore always show scatter. The regressions calculated using the two optional independent variables will not be identical and the difference will increase with increasing scatter.

The variable that is assumed to be most affected by the largest variability (scatter) is the time variable and it has to be considered as a dependent variable (random variable) in order to allow a correct statistical treatment of the data set in accordance with this method. However, for practical reasons, the industry prefers to present stress as a function of time as an independent variable.

Use of the methods [2]

This extrapolation method is designed to meet the following two requirements:

a) To estimate the lower prediction limit1) (at 97,5 % probability level) of the stress which a pipe made of the material under consideration is able to withstand for 50 years at an ambient temperature of 20 °C using water or air as the test environment.

b) To estimate the value of the lower prediction limit (at 97,5 % probability level) of the stress, either at different lifetimes or at different temperatures, or on occasion both.

There are several extrapolation models in existence, which have different numbers of terms. This SEM will use only models with two, three or four parameters.

Adding more terms could improve the fit but would also increase the uncertainty of the predictions.

The SEM describes a procedure for estimating the lower prediction limit (at 97,5 % probability level) whether a knee (which demonstrates the transition between type A and type B crack behaviour) is found or not (see Annex B).

The materials have to be tested in pipe form for the method to be applicable.

The final result of the SEM for a specific material is the lower prediction limit (at 97,5 % probability level) of the hydrostatic strength, expressed in terms of the hoop stress, at a given time and a given temperature.

1) In various ISO documents, the lower prediction limit (LPL) is referred to as the lower confidence limit (LCL), where LCL is the 97,5 % lower confidence limit for the mean hydrostatic strength.

The ISO document makes it very clear that there will always be scatter in the data due to non-ideal testing circumstances and material inhomogeneity. The ISO method requires the use of a statistically developed Lower Prediction Limit (LPL) in extrapolating from the empirically derived model. We will make extensive use of statistically derived prediction limits in the development of our RPM based model for FFS determination.

Discussion of Coleman and Bragaw Papers

Bragaw [3, 4] developed the RPM as a tool for predicting the service life of polyethylene pipes and fittings. Bragaw references an earlier work by Coleman [5] in which Coleman noted that the ultimate strain at break of polymeric filaments at a given temperature was invariant with regard to the rate of loading. This suggested to Coleman that a simple superposition principle could be developed for calculating the time to break by creep failure under an arbitrary loading history. The model that he selected was that of rate process theory, where the number of jumps away from their nearest neighbors that elements of the polymer chain take, is the process whose rate we are measuring. In other words, the RPM fundamentally measures the rate of creep of a polymer unit under stress at a given temperature. Coleman developed the mathematical model, performed experiments with multiple fiber constructions, and found that the time to failure is adequately described by a rate process model as shown in **Figure 1-2**.

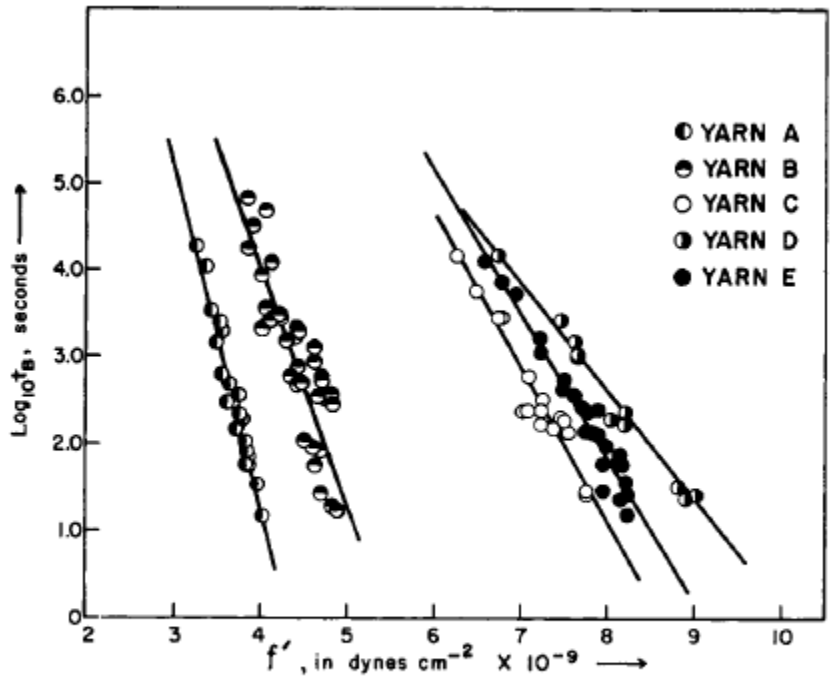


Fig. 1. Experimental plot of the logarithm of lifetime in seconds *versus* dead load in dynes cm.$^{-2}$. Stresses are calculated using the cross-sectional area of the unloaded yarn.

Figure 1-2. Linear relationship of $\log_{10}$(time to failure) as a function of stress [5]

Bragaw built upon Coleman's model and developed the RPM model as described by Kurdziel and Palermo [1]. Bragaw notes a key point that is highlighted in the excerpt from his 1982 paper in the insert below: **to correctly model the time to failure we need to know the true stress in the volume in which the failure is propagating through the pipe wall**. He points out the difficulties of determining this stress at the time of publication. At the present time, we have many more tools at our disposal to address the true stress problem and in this project, we have devoted considerable effort to developing Stress Intensification Factors (SIF) that can be used as a linear multiplier for hoop stress in a RPM equation. These methods will be described in detail in subsequent sections, but at this point it is sufficient to note that it is possible to calculate a SIF, or stress riser, that allows us to use a single reference RPM model, together with this SIF, to perform FFS calculations. A SIF (stress riser) of "1" describes a situation where the nominal hoop stress adequately describes the stress state of the pipe. We will use this SIF in the illustrative figures that follow.

Following Coleman(7), rate process theory describes the rate at which a rupture process proceeds by the equation

$$\frac{dc}{dt} = \frac{KT}{\eta} e^{-\left[\frac{\varepsilon-\beta\sigma/2}{KT}\right]} \quad (1)$$

where $\frac{dc}{dt}$ is fracture rate, t is time, T is absolute temperature, σ is true stress at activated sites, ε is activation energy, and the rest are constants. Coleman simplified this expression to find

$$\log t = A_0 T^{-1} + A_1 T^{-1}\sigma \quad (2)$$

Two difficulties arise when theoretical Equation 2 is applied to analysis of practical test data obtained on pressurized pipe, fittings, and joints. First, the stress σ in Equations 1 and 2 is the true stress in the activated volume near the crack tip. Because overall geometry is very complex, crack tip radius is unknown (and varies during propagation), and crack length varies, true stress σ cannot be calculated from system test pressure. Further, σ must vary with time under load even though pressure is maintained constant. Second, Equations 1 and 2 describe only a single activated process, and many competing processes may be present. The transition from the ductile failure mechanism to the slit failure mechanism is an example of one of many possible mechanistic transitions. The experimenter must carefully examine specimen failure mode to avoid mixing failures having different mechanisms.

The indeterminacy of true crack tip stress has led us to explore the statistical fit of a number of mathematical models relating temperature and pressure to test or service life of piping components. All mathematical models were closely related to theoretical Equation 1.

Excerpt from Bragaw, C.G. *The Forecast of Polyethylene Pipe and Fitting Burst Life Using Rate Process Theory*. in *5th International Conference on Plastic Pipes, York UK*. 1982.

Reference RPM Data for Aldyl A

GTI has a database of over 1400 Aldyl-A data points collected over the past decade. This data set includes about 400 reference data points generated by Palermo, working with Bragaw and others, at the DuPont company in their investigation and characterization of the Low Ductile Inner Wall (LDIW) condition that was known to cause premature Slow Crack Growth (SCG) failures of Aldyl A pipe. GTI also has several hundred well documented pipe and fitting failure points generated through long term testing of pipe exhumed from multiple areas after 40-45 years in service. This recent utility data set is consistent with the full data set that spans two decades of testing, indicating that the Aldyl-A pipe in the utility systems is no better than, and no worse than the piping systems evaluated to date.

To illustrate this point a reference temperature of 15°C (59F) has been selected, as it is the upper bound of the ground temperatures in many gas distribution systems. The distribution of the utility data relative to the reference data is the same for all temperatures, but the absolute stress versus time to failure values change from temperature to temperature.
A stress riser of one (1) is used in the following example plots to capture the behavior of pipe subjected to internal pressure alone.

In the charts that follow the diagonal lines depict the performance characteristics of the DuPont control material. The control material is material that performed as originally designed and specified in the official listing of the material. The solid black line is the mean, or nominal performance of the material at the reference temperature for the chart.

The additional lines depict prediction limits at various levels of confidence expressed in natural frequencies of the proportion of data points, sampled from the control pipe, that would be expected to fall below each line (e.g., 1 in 10, 1 in 100 etc.) for the lower prediction bounds, and 9 in 10, 99 in100 etc. for the upper prediction bounds.

The bands between the diagonal lines can be used to assign relative rankings (relative to the expected lifetime at the given stress and temperature) to the points falling in the various bands.

The open grey circles are all of the reference data GTI has collected over the years.
The colored symbols depict various DuPont and Utility Data Sets.

DuPont LDIW reference data was generated by DuPont using pipe known to exhibit the LDIW condition as defined by DuPont.

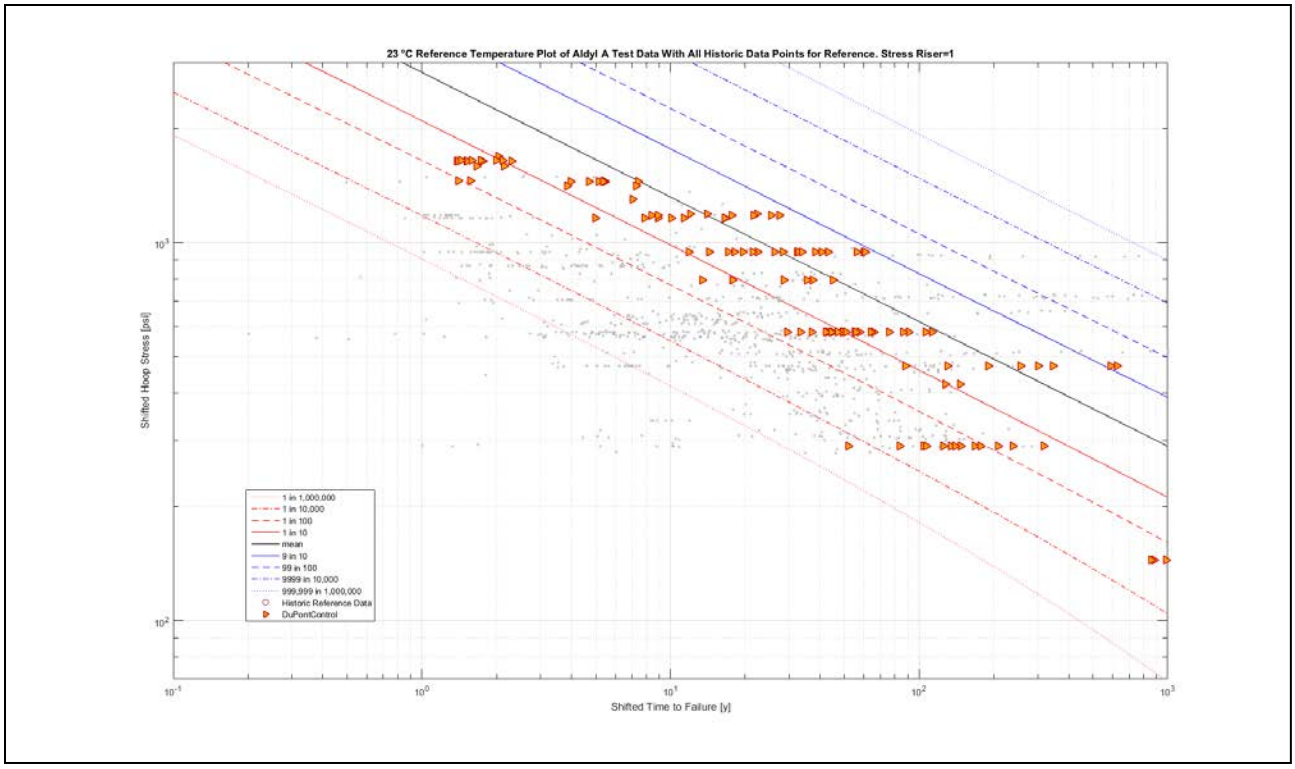


Figure 1-3 shows the control data developed by DuPont that will be used to develop the reference RPM model that forms the basis of the FFS calculation method developed in this project. **Figure 1-4** shows the DuPont LDIW data set together with utility data from exhumed pipe relative to the control data set.

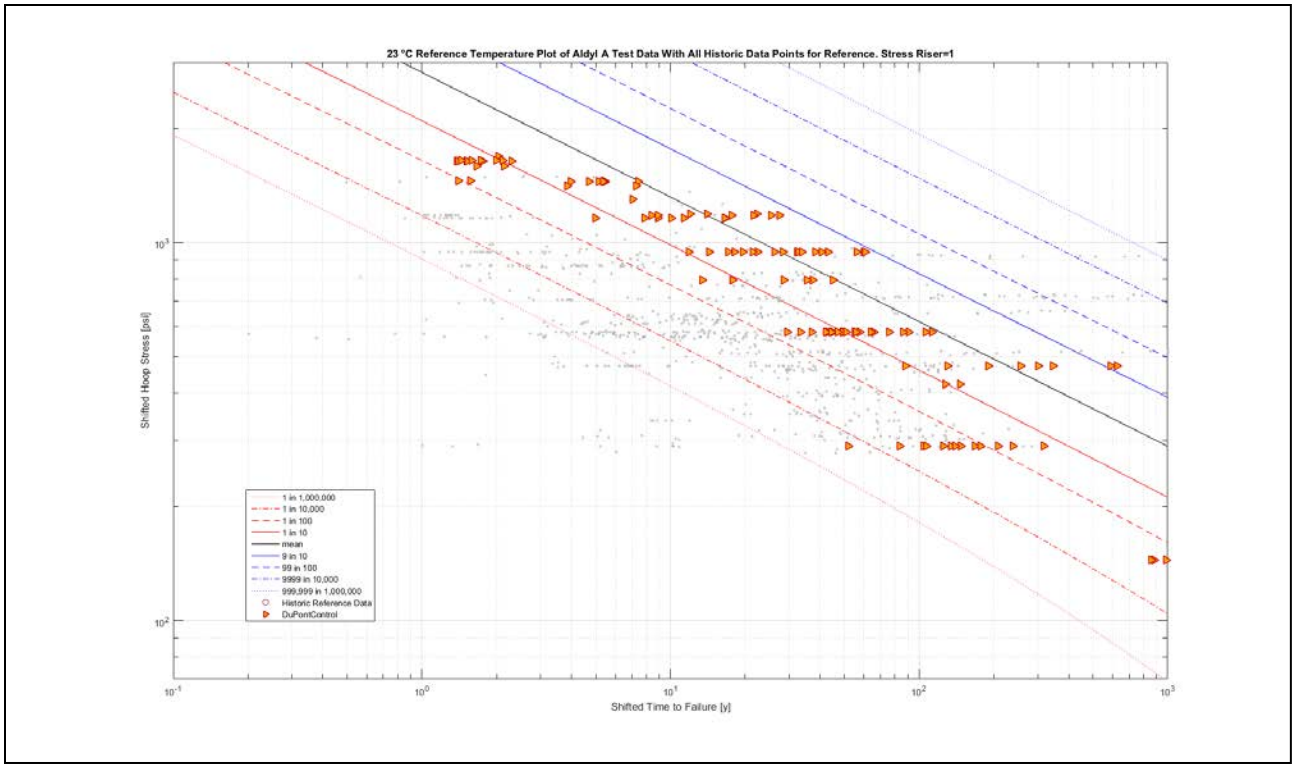


Figure 1-3. DuPont Control Aldyl-A at 23°C (73.4°F) Reference Temperature

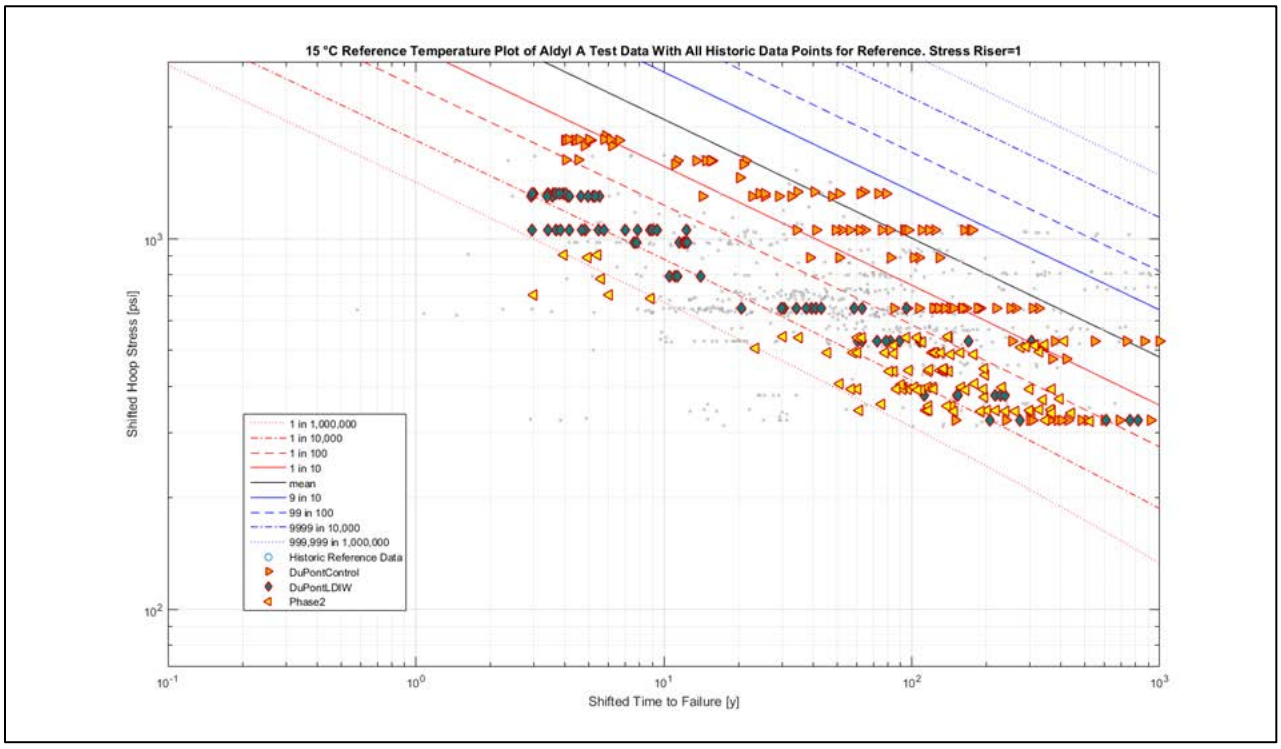


Figure 1-4. Data Set 5 Points at 15°C Showing Correspondence with DuPont LDIW Data

Three simple methods are available to us to assign easily-understood relative rankings to the individual data points generated via RPM testing

Material Ranking

The "Material Ranking" is based on the prediction limit bands calculated from the RPM model obtained from running multiple linear regression on the DuPont control data set according to ISO 9080 [2].

4.1 General model for the regression analysis according to ISO 9080

The general 4-parameter model used in ISO 9080 is the following:

$$Log(t) = C_1 + C_2 \cdot \frac{1}{T} + C_3 \cdot Log(\sigma) + C_4 \cdot \frac{Log(\sigma)}{T} + e$$

where

C_1 to C_4	*parameters used in this model*	
t	*time to failure*	*[h]*
T	*Temperature*	*[K]*
σ	*Hoop stress*	*[MPa]*
e	*error variable*	*Laplace-Gaussian distribution, with zero mean and constant variance (the errors are assumed to be independent)*

The 4-parameter model shall be reduced to a 3-parameter model if the probability level of C_3 is greater than 0.05. i.e. $C_3 = 0$.

Figure 1-5. The General RPM Model as Described by ISO 9080

Table 1-1. DuPont Control Model Parameters for 3 Parameter ISO 9080 Model

Parameters	C1	C2	C4	R^2	R^2_{adj}	n	p	σ^2
Value	-17.6172	9485.337	-898.536	0.898	0.896	122	3	0.080944
Standard Error	0.703265	296.4989	33.90417					

Covariance Matrix		
6.110133978	-2487.297563	152.8993481
-2487.297563	1086051.522	-89681.8456
152.8993481	-89681.8456	14201.01238

The material model fully described in **Table 1-1** enables us to calculate prediction intervals for any desired level of confidence using standard statistical methods as described by Montgomery [6] p468 where the prediction of new observations is discussed:

$$\hat{y}_0 - t_{\frac{\alpha}{2},n-p}\sqrt{\hat{\sigma}^2(1 + x_0'(X'X)^{-1}x_0} \leq Y_0 \leq \hat{y}_0 + t_{\frac{\alpha}{2},n-p}\sqrt{\hat{\sigma}^2(1 + x_0'(X'X)^{-1}x_0}$$ **Equation 1-1**

Y_0 – response variable, in this case the predicted lifetime in hours
X_0 – the input vector [1 1/T Log(s)]. Note: Log refers to <u>log base 10</u> in this equation
$(\mathbf{X'X})^{-1}$ – is the covariance matrix given in **Table 1-1**
T – absolute ambient temperature in Kelvin
s – hoop stress in the pipe at operating pressure in psi
σ^2 – the variance of the data set
α - the desired confidence level e.g. 0.05 for a 95% two-sided confidence interval
n – the number of samples in the data set
p – the number of parameters being estimated, 3 in this instance
t – the t statistic for the desired confidence and degrees of freedom in the data set

Equation 1-1 was used to generate the diagonal lower and upper prediction limit lines shown in the plots above. The confidence levels were set to reflect natural frequencies of 1 in 10, 1 in 100, 1 in 10,000 and 1 in 1,000,000 for the lower prediction limits and 9 in 10, 99 in 100, 9,999 in 10,000 and 999,999 in 1,000,000 for the upper prediction limits. Natural frequencies are easier to grasp in this context – they reflect the number of future observations we would expect to fall below each of the prediction limit lines.

The analysis described above is very formal and reflects the confidence we have in the DuPont reference data set. The data set is good as can be seen in low variance and good R^2 of the model.

The prediction bands, Figure 1-6. Relative Ranking Bands for Material RPM Performance Relative to DuPont Reference Data**Figure 1-6,** are appropriate for highlighting how the data sets we develop from samples extracted from the field perform in an absolute sense relative to the DuPont data.

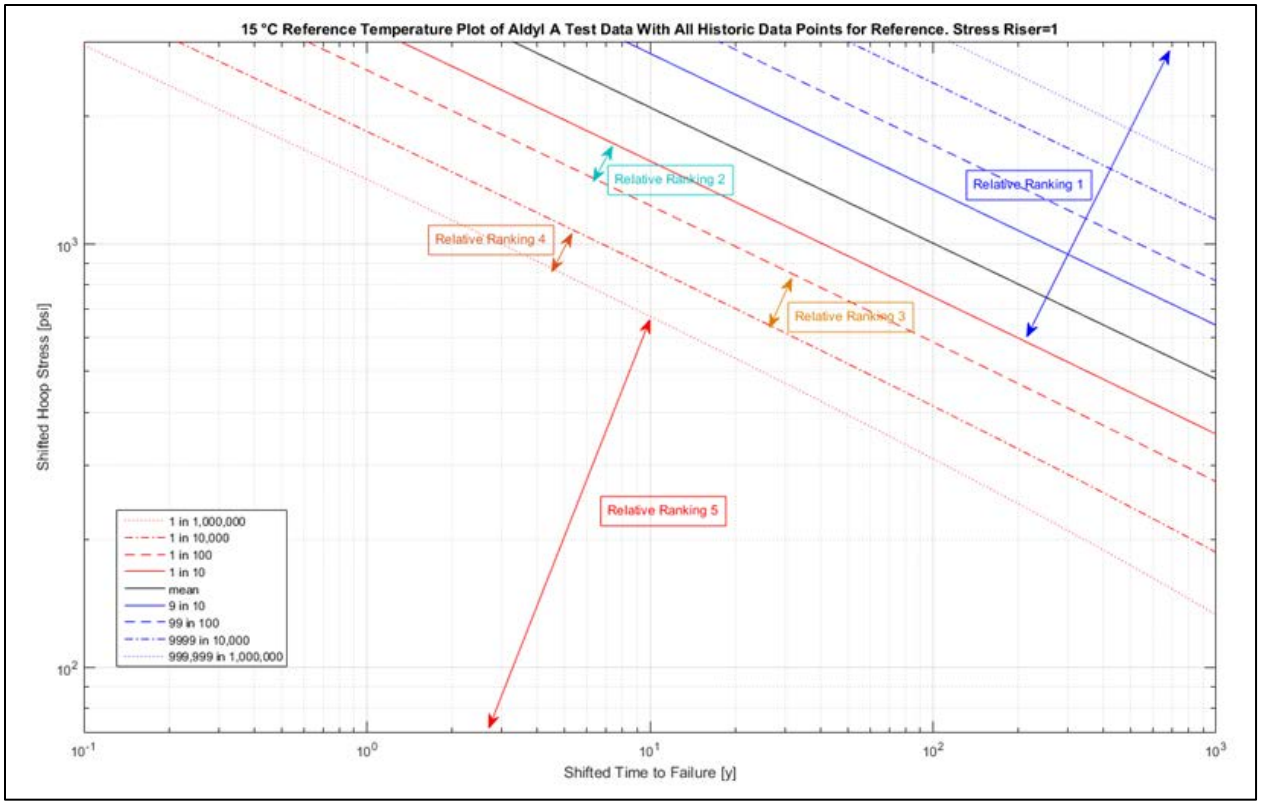


Figure 1-6. Relative Ranking Bands for Material RPM Performance Relative to DuPont Reference Data

There is a drawback in looking only at this relative ranking as data can fall in a band that indicates that the buried pipe is performing well below the reference performance. This may be true in an absolute sense, but does not address the operational implications of this below-par material performance.

Two additional reference rankings can be used to address the operational implications:

1. The hoop stress at which the specimen failed
2. The time at which the specimen failed

These two additional rankings can be scaled relative to the operating stress of the pipeline and the desired residual pipe lifetime.

Failure Stress Ranking

The system in question operates at 45 psig that will cause a hoop stress in SDR 11 pipe of 225 psi. In this report the relative ranking levels were chosen to be 100 psi, 200 psi, 300 psi, and 400 psi for convenience. This choice is arbitrary and subjective and can be changed by subject matter expert consensus at any time. **Figure 1-7**

Failure Time Ranking

The failure time ranking levels were chosen to be 10 years, 20 years, 30 years, and 40 years. This choice is arbitrary and subjective and can be changed by subject matter expert consensus at any time. **Figure 1-8**

Operational Relative Ranking

The three rankings described above can be combined into a relative ranking that takes into account the material performance, system operating pressure and the desired residual lifetime of the pipe as follows:

$$\boldsymbol{Operational\ Ranking} = \sqrt[3]{\boldsymbol{Material\ Ranking} * \boldsymbol{Failure\ Stress\ Ranking} * \boldsymbol{Failure\ Time\ Ranking}} \qquad \textbf{Equation 1-2}$$

The relative ranking score obtained by applying **Equation 1-2** scales from 1 to 5. The component scores are integers, but the resultant operation ranking will be a real number.

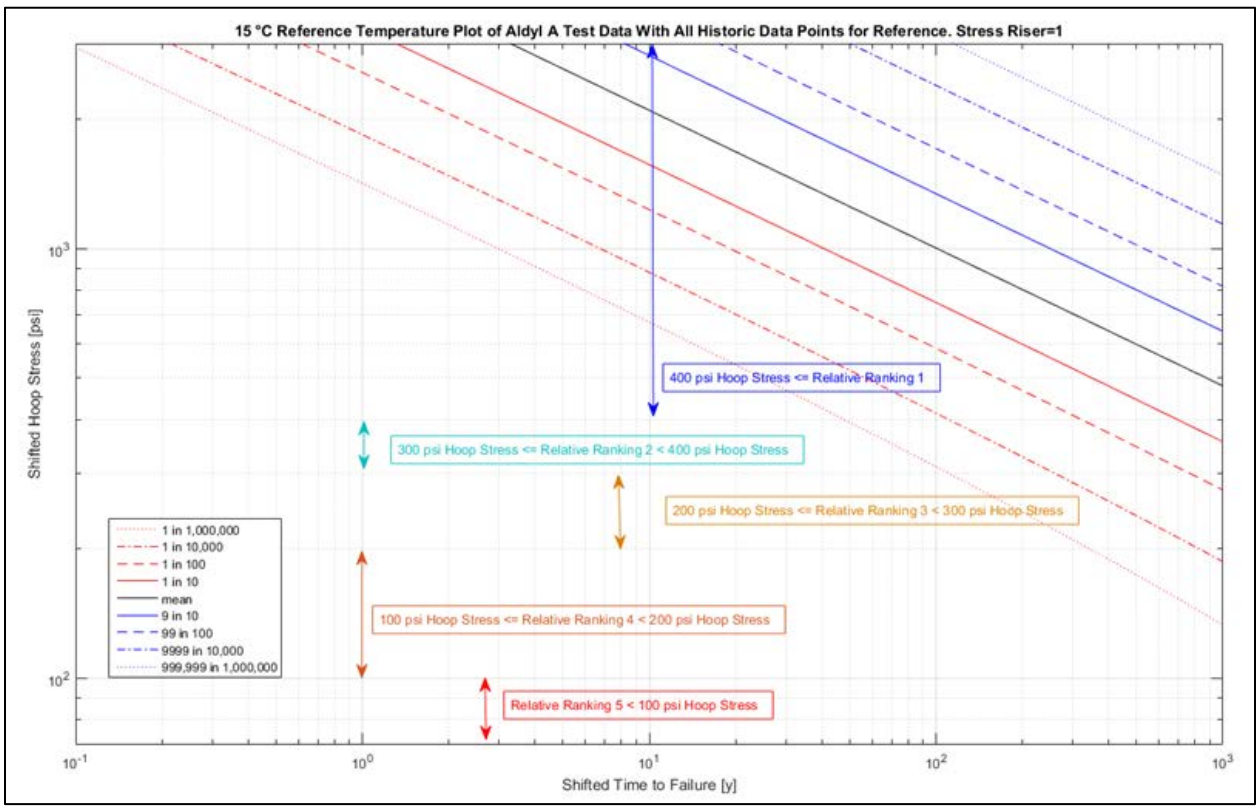


Figure 1-7. Relative Ranking Bands for Failure Stress Relative to Absolute Hoop Stress in psi

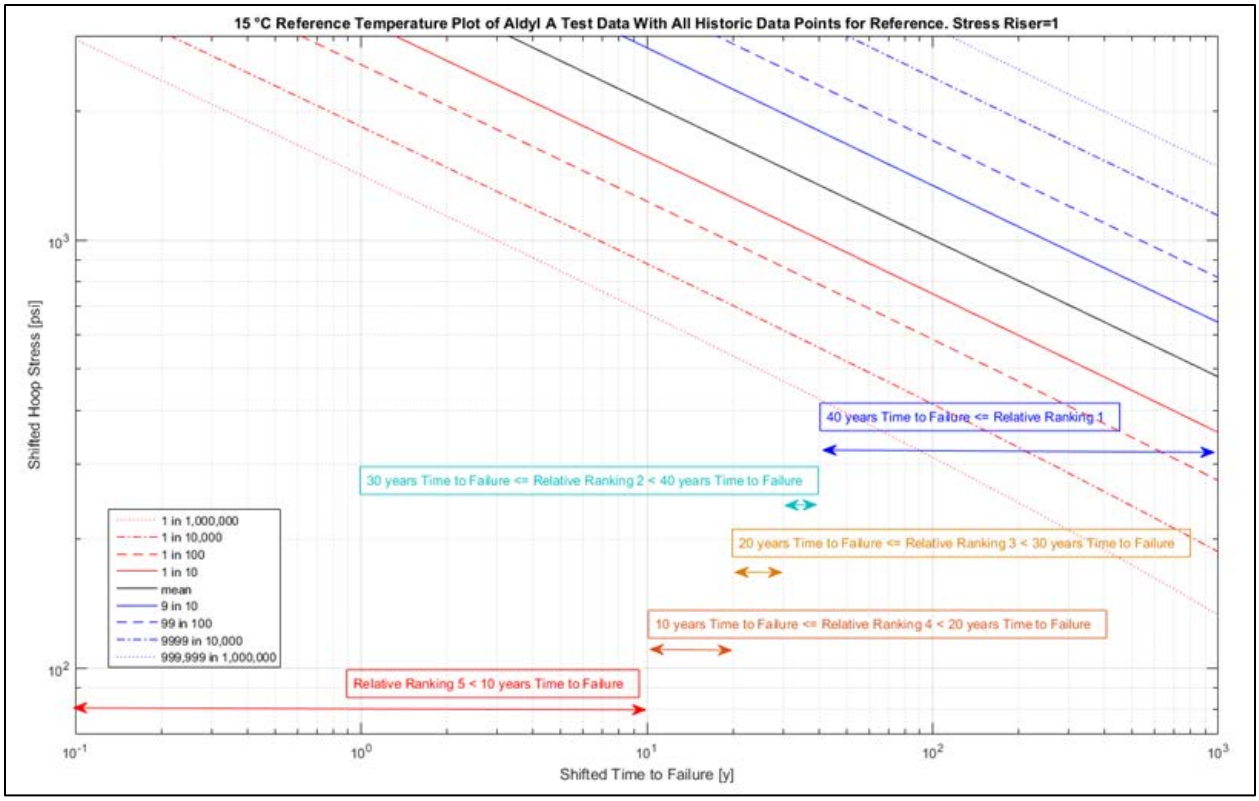


Figure 1-8. Relative Ranking Bands for Failure Time Relative to Absolute Time to Failure in years

Some typical output (Data Set 1, 1969 vintage Aldyl A, approximately 45 years in service) is presented in **Figure 1-9** to **Figure 1-12**. We can see that the materials RPM performance puts it in medium to high risk bands, however the test stresses translate to very high operational pressures at the reference temperature and all of the failure points are in a low operational risk band looking at operating stress alone. Looking at the projected times to failure we see that most of the points are projected to relatively low failure times, placing them in medium to very high risk looking at projected failure times alone. The composite operational risk is low to medium in the absence of fittings, squeeze-offs, or other installation conditions that could introduce a SIF. It is extremely important to have good knowledge of potential SIF to conduct a complete FFS evaluation. We will comprehensively address SIF in later section of the report.

The histograms are a simple count of how many individual data points from the long-term hydrostatic RPM testing fell into each ranking category for the data sub-set being analyzed. A histogram is presented for each of the three rankings described above for a single illustrative example: Data Set 1 (1969), 15°C ground temperature and a stress riser of 1 representing straight pipe under hydrostatic pressure with no fitting or installation condition induced stress risers.

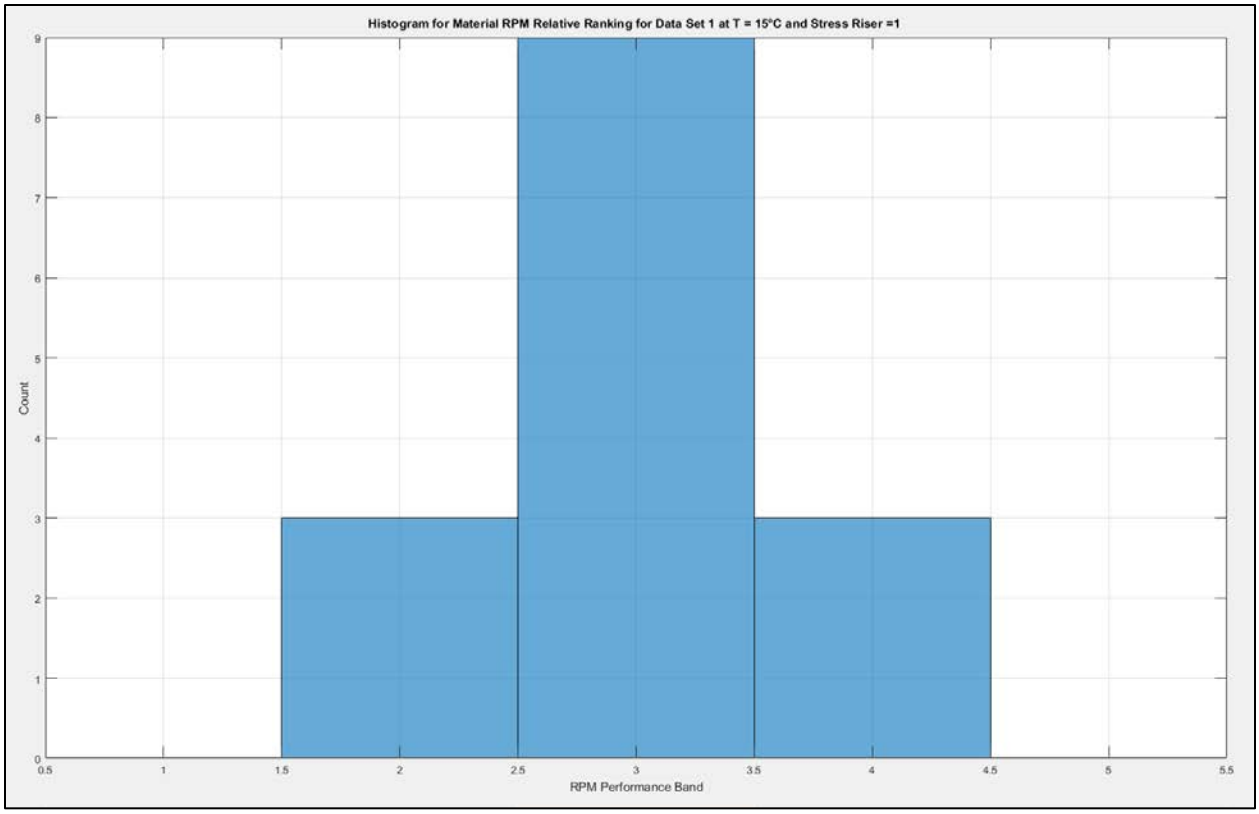


Figure 1-9. Material RPM Relative Ranking for Data Set 1 @ 15°C

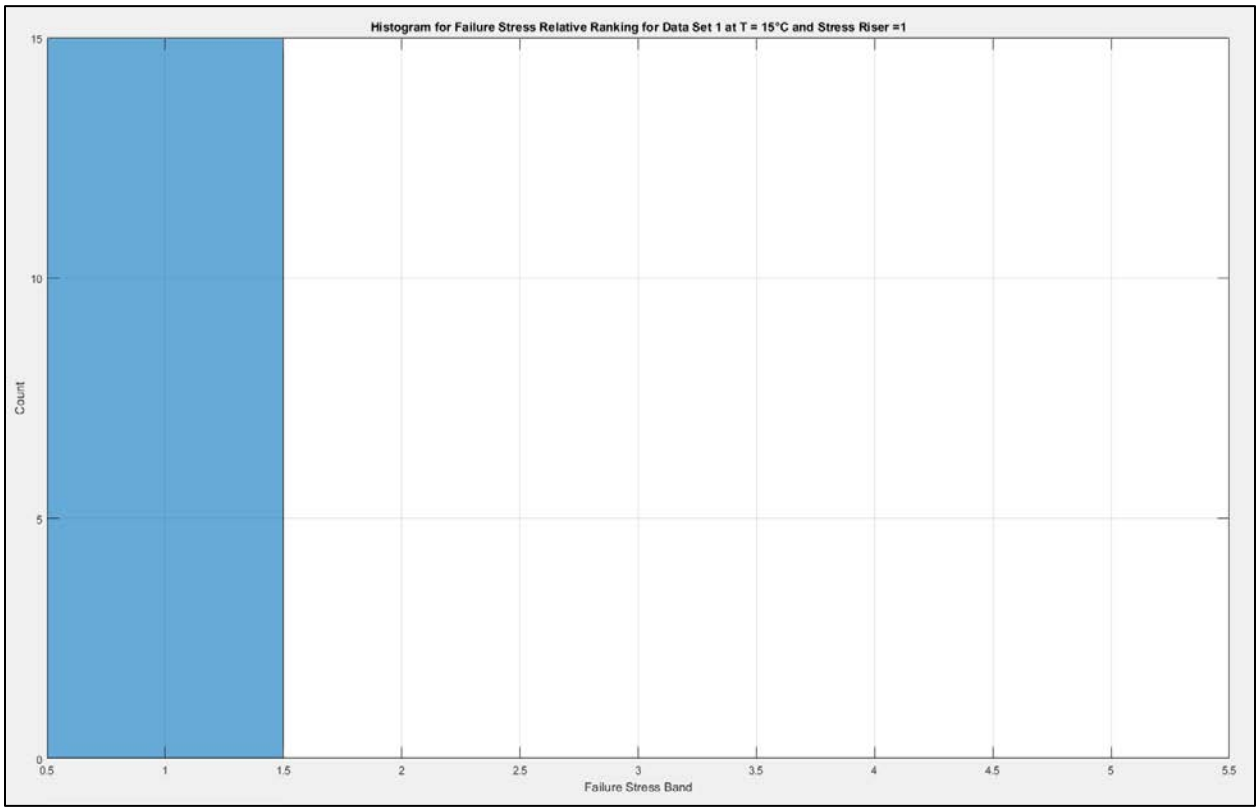


Figure 1-10. Failure Stress Relative Ranking for Data Set 1 @ 15°C

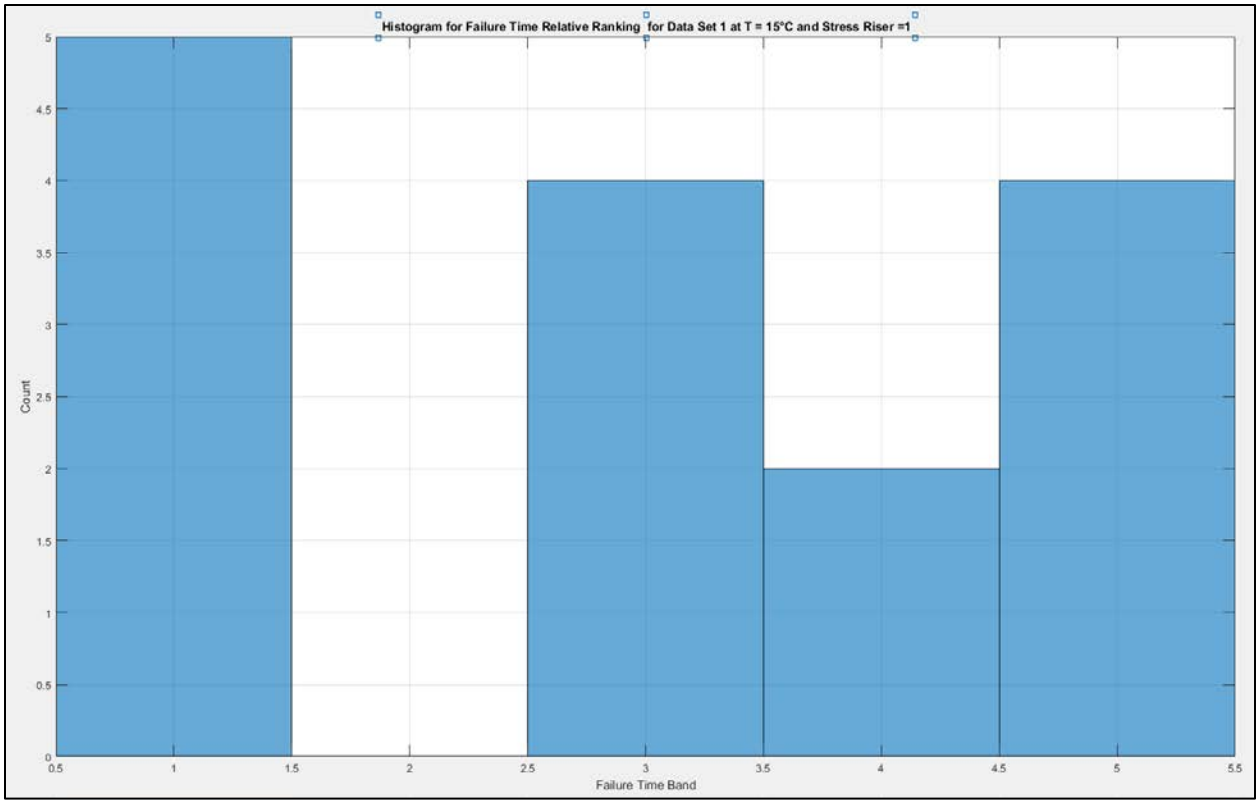


Figure 1-11. Failure Time Relative Ranking for Data Set 1 @ 15°C

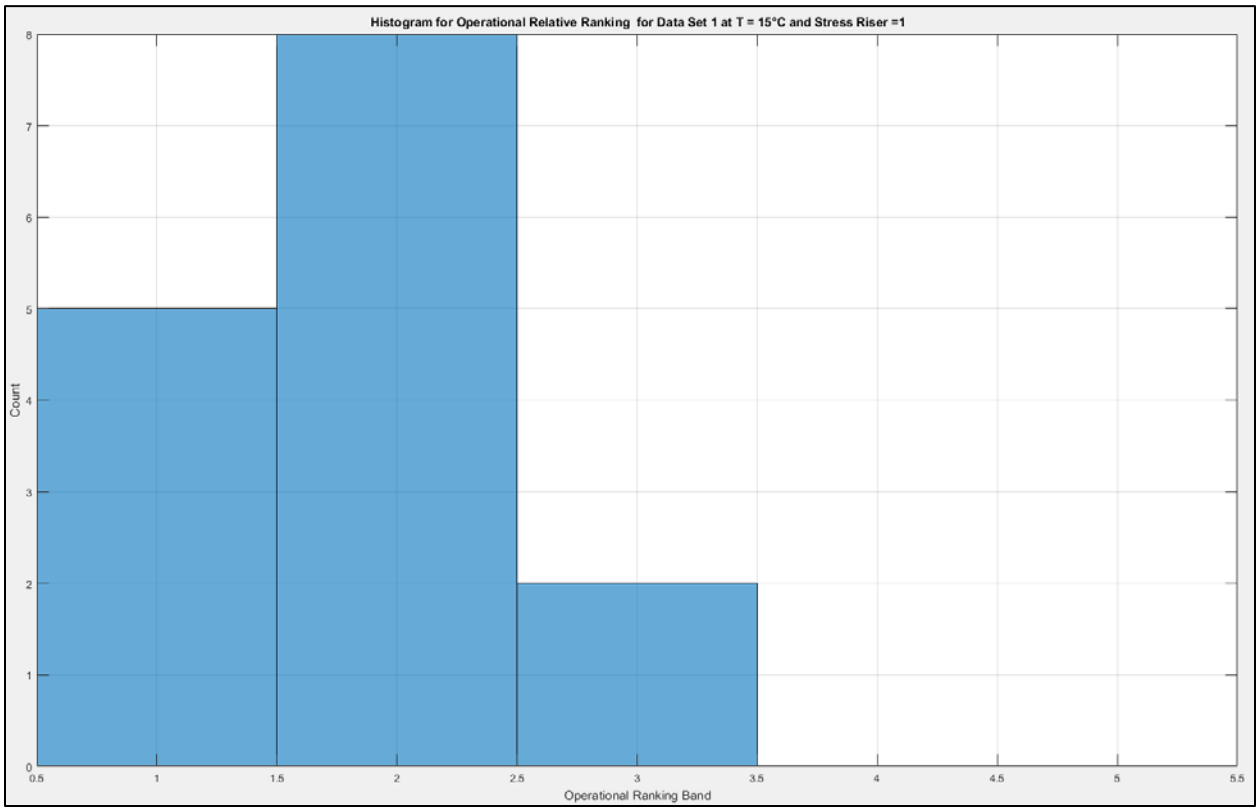


Figure 1-12. Operational Relative Ranking for Data Set 1 @ 15°C

The Impact of Pipe Inner Wall Surface Condition on Ranking Scores

During the investigations carried out under the Operations Technology Development (OTD) funded project 2.8.d (GTI 20649) that developed risk models for Aldyl A gas distribution piping, it was noted that certain surface features visible in Cross-Polarized Light Microscopy (CPLM) and Scanning Electron Microscopy (SEM) appeared to correlate to times to failure in the RPM testing of pipe specimens. A categorical logistic regression model yielded surprisingly good predictive power as shown in **Figure 1-13**. The results of this study were presented at Plastics Pipes XVII in Chicago, IL September 2014 [7]. In this paper the term "Risk" is used out of context relative to risk assessment where risk is probability of occurrence multiplied by the consequence of the probable event if it occurs. Here the usage of *risk* is: "someone or something that creates or suggests a hazard" e.g. "There is a *risk* of liver damage with this medication", "Pipes with these surface features present have an increased *risk* of reduced lifetime expectancies due to the lower resistance to damage associated with the presence of these features.".

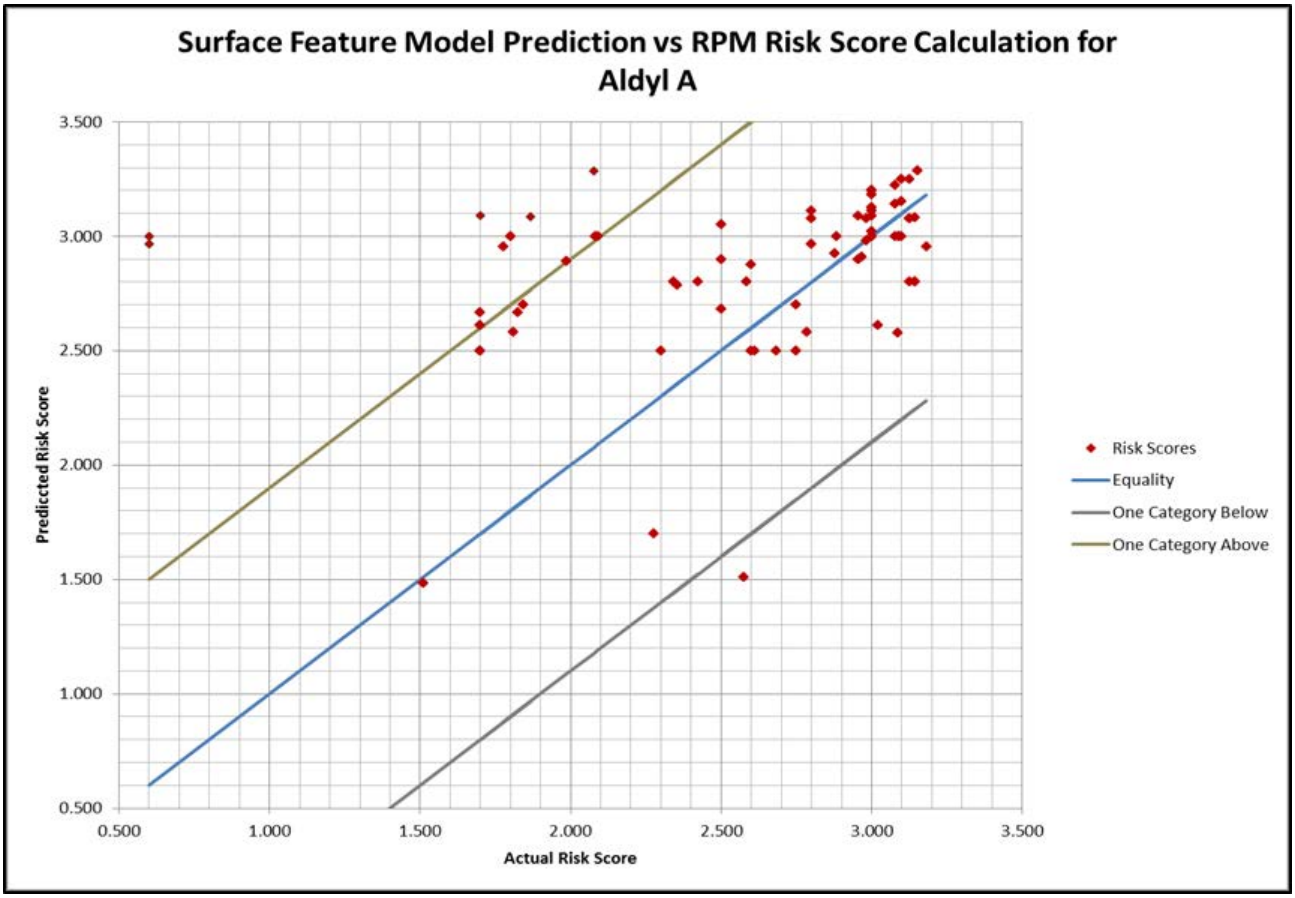


Figure 1-13. Initial Surface Feature Model for Predicting "Risk" Scores

This kind of model is extremely useful for pipeline operators in that it can help narrow down the conditional probability estimates of low lifetime expectancy given the presence of these

features. **Table 1-2** shows visually how the presence of these features correlates with "Risk" ranking calculated from the DuPont control RPM model and the DuPont LDIW RPM model.

Table 1-2. Data for Initial Surface Feature Model

Risk					CPLM	Risk					CPLM
Rank	Dimple	Micro crack	Crystal	Fiber (rod)		Rank	Dimple	Micro crack	Crystal	Fiber (rod)	
0.6	1	0	1	0	15	2.8	0	0	1	1	10
0.6	1	1	1	1	10	2.8	0	0	0	0	10
1.4	0	0	0	0	20	2.8	0	0	1	1	10
1.7	1	1	0	0	10	2.8	1	1	1	1	10
1.7	1	1	0	1	10	2.8	0	0	1	1	10
1.7	1	1	0	1	10	2.8	1	1	1	0	15
1.7	1	1	0	0	10	2.8	1	1	1	0	20
1.7	0	0	0	0	10	2.8	1	0	1	1	10
1.7	0	1	0	0	10	2.8	1	0	1	1	15
1.7	0	0	0	0	10	2.8	1	1	1	1	15
1.7	0	0	1	0	10	2.8	1	0	1	1	15
1.7	0	0	1	0	20	3.0	1	1	1	1	20
1.7	0	0	1	1	20	3.0	1	1	1	1	20
1.7	0	0	1	1	10	3.0	1	1	1	1	20
1.7	0	0	1	0	10	3.0	0	0	1	0	20
2.0	1	1	1	0	25	3.0	0	0	1	1	20
2.0	1	1	1	1	30	3.0	1	1	1	0	25
2.0	0	1	1	0	20	3.0	1	1	1	0	25
2.2	0	0	0	0	15	3.0	1	1	1	1	25
2.2	1	1	1	1	10	3.0	1	1	1	1	20
2.2	1	1	0	1	10	3.0	1	0	1	1	20
2.2	1	0	0	0	10	3.0	1	1	0	0	25
2.2	1	1	1	1	10	3.0	1	1	1	0	25
2.5	1	1	0	0	10	3.0	1	1	1	0	10
2.5	1	1	1	1	10	3.0	1	1	1	0	25
2.5	1	1	1	0	10	3.0	1	0	1	0	10
2.5	0	0	0	0	10	3.0	1	1	1	1	30
2.5	0	0	1	0	10	3.0	1	1	0	1	25
2.5	1	1	1	1	10	3.0	1	1	1	0	25
2.5	0	0	0	0	10	3.0	1	1	0	1	25
2.5	0	0	0	0	10	3.0	1	1	1	0	10
2.5	1	1	0	0	10	3.0	1	1	1	0	25
2.5	0	0	0	0	20	3.0	1	1	1	1	30
2.5	0	0	0	0	10	3.0	1	1	1	0	30
2.5	0	0	0	0	10	3.0	1	1	1	0	30
2.8	0	0	1	0	10	3.0	1	1	1	1	10
2.8	0	1	0	0	10						

Figure 1-14 and **Figure 1-15** show examples of the surface features.

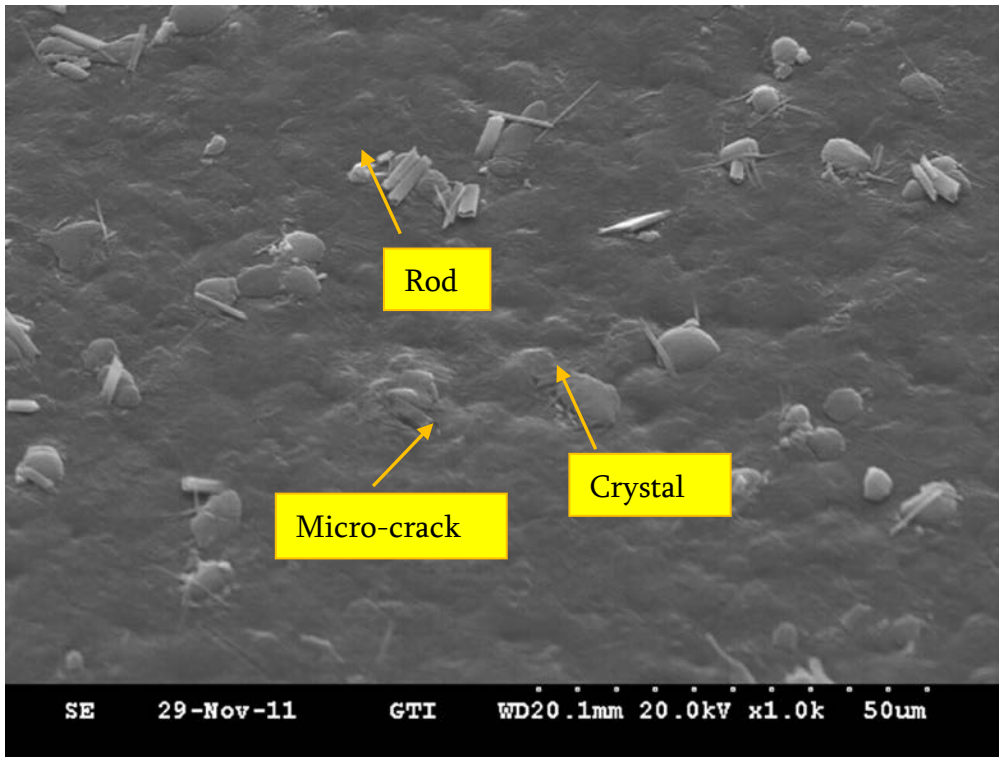


Figure 1-14. SEM image showing Crystals, Rods and Micro-crack

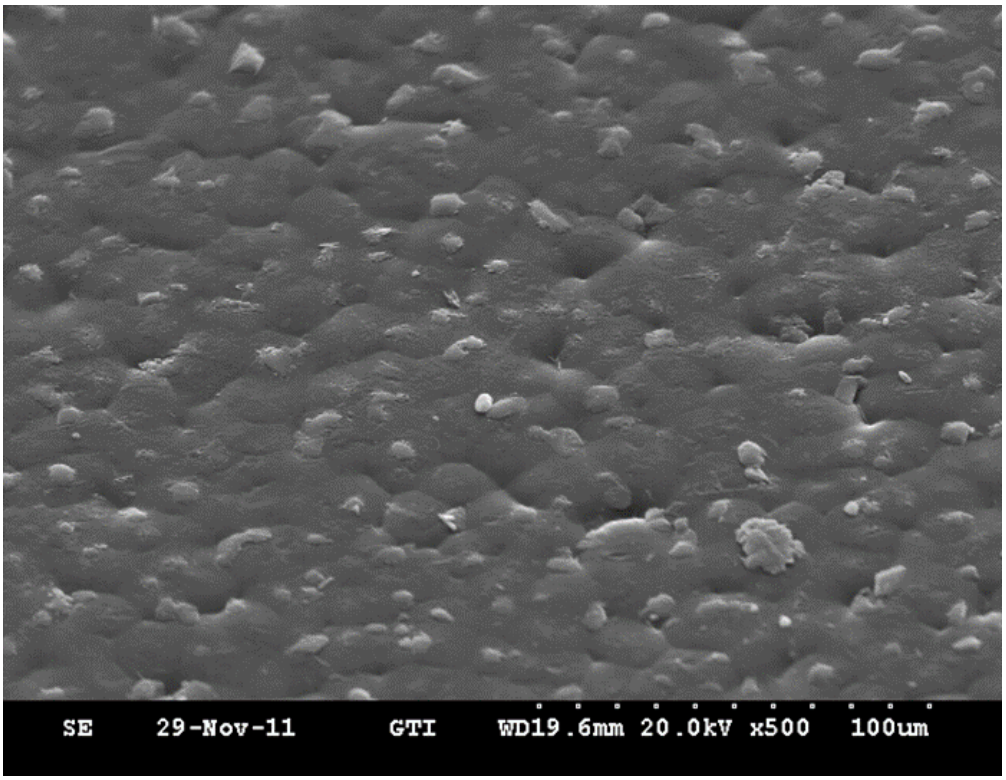


Figure 1-15. SEM Image Showing Dimples

Table 1-3 shows the data for an improved surface feature model where the initial surface feature model was improved by including relative rankings derived from Oxidation Induction Time (OIT) from thermal testing and Carbonyl Index (CI) from Fourier Transform Infrared (FTIR) testing.

Table 1-3. Data for Expanded Surface Feature and Thermal Characteristics Model

LIMSNo	LTHSId	SampleID	OIT SampleID	Dimple	Micro crack	Crystal	Fiber (rod)	CPLM	OIT Risk	FTIR Risk	Risk Score
111174-001	452	E1A	E1A	1	1	1	1	10	1	1	0.6
111174-002	455	E1B	E1AB	1	0	1	0	15	1	1	0.6
111174-006	456	E5B	E5BC	1	1	0	1	10	2	1	1.1
111174-012	451	E10B	E10A	1	1	0	0	10	3		1.7
111174-014	457	E12	E12	1	1	0	1	10	2	1	1.7
111174-015	458	E47	E47	1	1	0	0	10	1		1.7
111174-018	461	E200C	E200C	1	1	1	0	25	3	1	1.7
111174-019	462	E207A	E207AB	1	1	1	1	30	3	1	1.7
111174-029	472	E211C	E211CD	0	0	0	0	10	3	1	1.7
111174-030	473	E213A	E213A	0	1	0	0	10	3	1	1.7
111174-031	474	E213B	E213B	0	0	0	0	10	3	1	1.7
111174-032	475	E217A	E217AB	0	0	1	0	20	3		1.7
111174-034	477	E217C	E217BC	0	0	1	0		3	1	1.7
111174-035	478	E217D	E217CD	0	0	1	1	20	3		1.7
111174-036	479	E224A	E224AB	0	0	1	1	10	3	1	1.7
111174-037	480	E224B	E224B	0	0	1	0	10	3	1	1.7
111174-039	482	E224D	E224D	0	0	0	0		3		1.7
111174-041	484	E226B	E226B	0	1	1	0	20	3	1	2
111174-042	485	E230A	E230A	0	0	1	0	10	3	1	2
111174-043	486	E230B	E230BC	0	0	0	0	15	3	1	2.2
111174-044	487	E230C	E230CD	1	1	0	1	10	2	1	2.2
111428-001	489	E401A	E401A	1	1	0	1		2		2.2
111428-002	490	E401B	E401BC	1	1	0	0	10	1		2.2
111428-005	491	E401E	E401E	1	1	1	1	10	1		2.2
111428-007	492	E402B	E402AB	1	1	1	0	10	1		2.2
111428-009	493	E402D	E402CD	0	0	1	0	10	1		2.2
111428-010	494	E402E	E402DE	0	0	0	0	10	3		2.2
111428-012	496	E403B	E403CD	1	0	1	1	20	3		2.2
111428-015	497	E403E	E403DE	1	1	0	1	25	3	1	2.2
111428-016	498	E404A	E404AB	1	1	1	0	10	3	1	2.2
111428-019	499	E404D	E404DE	1	1	1	1	30	3	2	2.2
111428-021	500	E405A	E405AB	1	1	1	0	30	3	1	2.2
111428-023	502	E405C	E405CD	0	0	0	0	20	3	1	2.2
111428-025	504	E405E	E405DE	0	0	1	0	10	3	1	2.2
111428-029	505	E406D	E406D	1	1	1	1	10	3	1	2.5
111428-032	506	E407C	E407CD	1	1	1	1	10	1	1	2.5
111428-034	507	E407E	E407E	1	1	0	0		2		2.5
111428-035	508	E408A	E408A	0	0	1	1	10	3		2.5
111428-041	509	E410A	E410AB	0	0	1	1	10	3		2.5
111428-042	510	E410B	E410BC	1	1	1	1	15	3	1	2.5
111428-044	511	E410D	E410DE	1	1	1	1	20	3		2.5
111428-046	512	E411A	E411A	0	0	1	1	20	3		2.5
111428-048	513	E411C	E411CD	1	1	1	0	10	3	1	2.5

LIMSNo	LTHSId	SampleID	OIT SampleID	Dimple	Micro crack	Crystal	Fiber (rod)	CPLM	OIT Risk	FTIR Risk	Risk Score
111428-048	513	E411C	E411CD	1	1	1	0	10	3	1	2.5
111428-053	514	E412C	E412BC	1	1	1	0		3	1	2.5
111428-054	515	E412D	E412DE	1	1	0	1	25	3		2.5
111429-001	516	E413A	E413A	1	1	1	0	25	3		2.5
111429-007	518	E414A	E414A	1	1	1	0		3	1	2.5
111429-012	519	E415A	E415A	1	1	1	0	30	3	2	2.5
111429-018	520	E416B	E416A	0	0	0	0	10	3	1	2.5
111429-024	521	E417C	E417CD	1	0	0	0	10	2		2.8
111429-026	522	E417E	E417E	0	0	0	0	10	3	1	2.8
111429-027	523	E418A	E418AB	0	1	0	0	10	3		2.8
111429-031	525	E418E	E418DE	0	0	1	1	10	3		2.8
111429-033	526	E419B	E419A	1	1	1	0	20	3	1	2.8
111429-035	527	E419D	E419CD	1	0	1	1	15	3	1	2.8
111429-040	528	E420D	E420E	1	0	1	1	15	3		2.8
111429-042	529	E421A	E421AB	1	1	1	1	20	3	1	2.8
111429-045	530	E421D	E421DE	0	0	1	0	20	3		2.8
111429-048	531	E422B	E422AB	1	1	1	0	25	3		2.8
111429-049	532	E422C	E422CD	1	1	1	0	25	3	1	2.8
111429-050	533	E422D	E422E	1	0	1	0	10	3	1	2.8
131119-012	543	E19	E2	1	1	1	1	30	3	1	2.8
131119-016	547	E21B	E21B	1	1	1	0	25	3	1	2.8
131119-018	549	E331A	E331AB	0	0	0	0		3	1	2.8
131119-019	550	E331C	E331CD	1	1	1	1	10	2		3
131119-021	552	E332A	E332BC	1	1	1	1	10	3	1	3
131119-022	553	E332D	E332CD	1	1	1	1	10	3		3
131119-023	554	E333A	E333AB	1	1	1	0	15	3	1	3
131119-024	555	E333D	E333DE	1	0	1	1	10	3		3
131119-025	556	E334A	E334BC	1	1	1	1	20	3	1	3
131119-026	557	E334D	E334D	1	1	1	0	25	3		3
131119-027	558	E335A	E335AB	1	1	1	1	25	3	1	3
131119-028	559	E335B	E335CD	1	1	1	1	20	3		3
131119-029	560	E336A	E336AB	1	1	0	0	25	3		3
131119-030	561	E336C	E336CD	1	1	1	0	25	3	1	3
131119-031	562	E337A	E337AB	0	0	0	0	20	3		3
131119-032	563	E337C	E337CD	1	1	1	0		3	3	3
131119-033	564	E337D	E337DE	0	0	0	0	10	3		3
131119-035	566	E338C	E338BC	0	0	0	0	10	3		3
131119-036	567	E338D	E338DE	1	1	0	0	10	3		3
131119-037	568	E338E	E338EF	0	0	0	0	10	3	1	3

Figure 1-16 shows the performance of the improved model graphically.

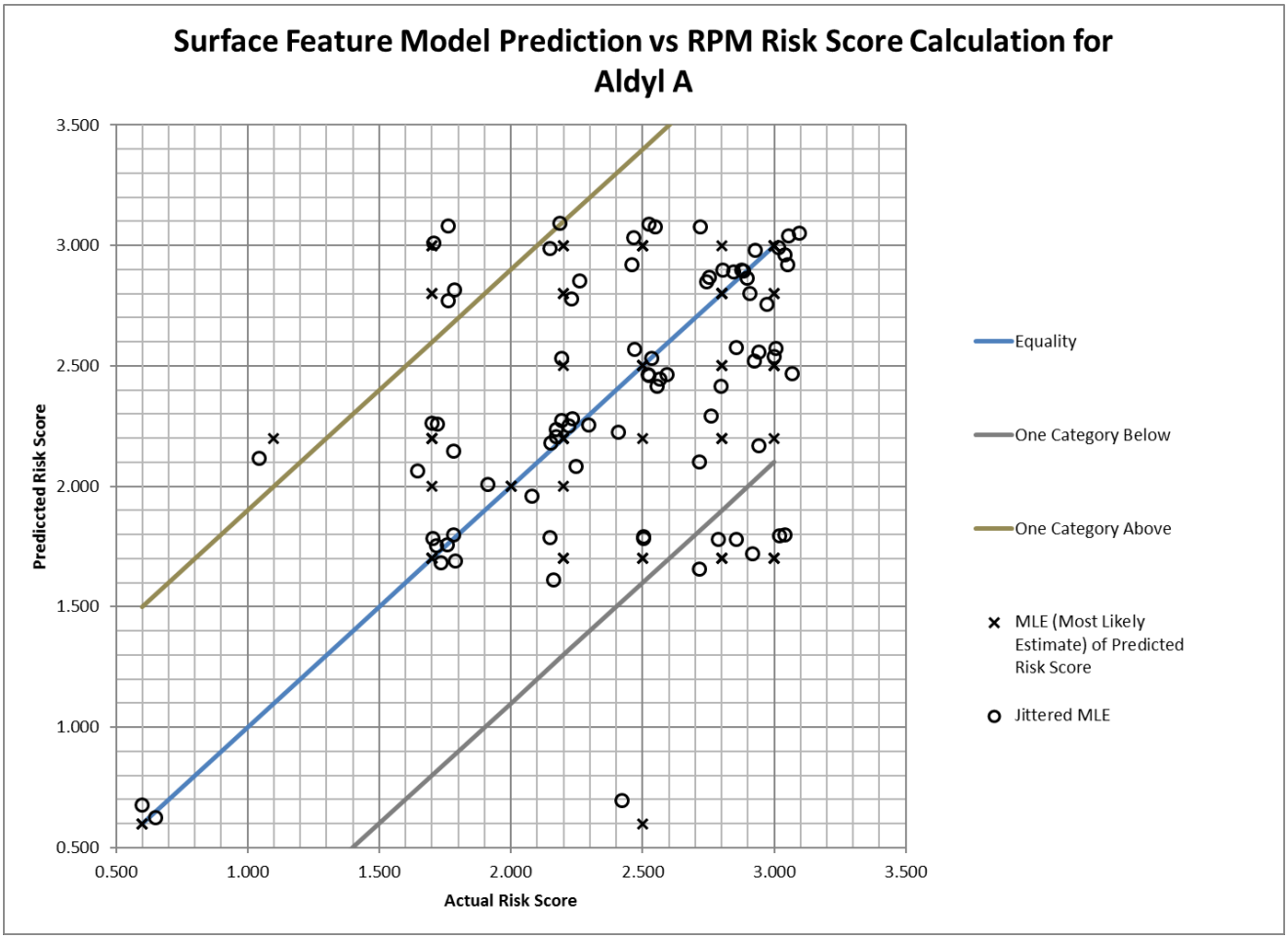


Figure 1-16. Revised Surface Feature Model

Surface Condition Ranking Model V2.0

The early surface correlation model was further refined for this project by rigorously documenting the surface condition and thermal characteristics of 16 long-term hydrostatic test specimens of Aldyl A pipe from a large set of exhumed pipe covering vintages from 1969 to 1974. The data set is presented in **Table 1-4**.

Table 1-4 – Improved Surface Condition Ranking Model: 14 correct predictions, 2 conservative errors; 87.5% success rate

Dimple	Micro Crack	Rod	Boundary Crystals	Surface Crystals	OIT	FTIR - CI		RPM Ranking		Predicted RPM Ranking
1	1	0	1	2	3	5		3	✓	3
1	1	1	1	1	4	4		4	✓	4
1	0	1	1	1	2	3		3	✓	3
0	1	0	1	1	5	1		2	✗	3
1	1	0	1	1	5	1		3	✓	3
1	1	0	1	3	5	1		2	✓	2
0	1	0	1	1	5	1		3	✓	3
1	1	0	4	1	5	1		2	✓	2
0	1	1	1	1	3	1		2	✓	2
0	1	0	1	1	2	1		2	✓	2
1	1	0	3	1	1	2		2	✗	3
1	1	1	4	5	2	4		3	✓	3
1	1	0	3	2	1	3		3	✓	3
1	0	0	3	1	1	3		3	✓	3
1	1	0	3	1	1	2		3	✓	3
1	1	0	5	2	1	2		2	✓	2

The inputs considered for the model were:

1. Binary for the presence, or absence of:
    a. dimples,
    b. micro-cracks, and
    c. rods.
2. Scaled for the presence of crystals on:
    a. The spherulitic boundaries, and
    b. Across the surface of the spherulites.
3. OIT results, and
4. FTIR results quantified via the Carbonyl Index (CI).

A method for counting the number of crystals per unit length, or area was developed and the count was normalized to a scale of 1 to 5 where 1 reflects the minimum number and 5 the highest number of crystals observed. The OIT and FTIR results were scaled in a similar manner.

A logistic regression model was developed for predicting the Material Ranking Categorization established from the RPM testing described above.

The regression model correctly predicts 14 of the 16 results, with the missed pair of points being conservatively over estimated for severity of ranking. The success rate of the model is 87.5% (14/16) – the results are shown graphically in **Figure 1-17**.

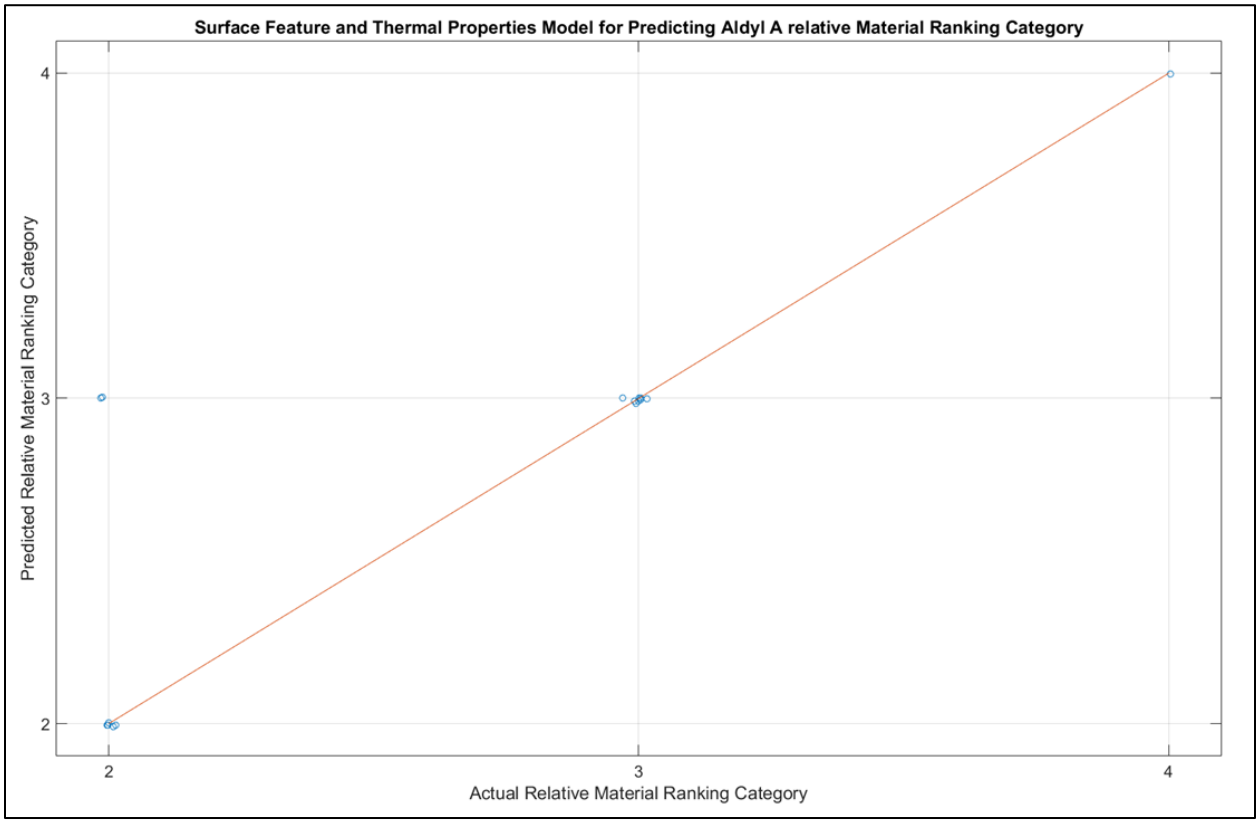


Figure 1-17. Performance of Surface Feature and Thermal Properties Model for Material Ranking Category Prediction

In **Figure 1-18** to **Figure 1-33** we show the probabilistic RPM performance band prediction from surface features and thermal characteristics of each specimen. We do this to emphasize that the prediction is probabilistic and that we should select the band with the highest probability (most likely estimate). We can see that for most data points the choice is clear cut, but for some there is close to a 50% likelihood that the point falls in each of two adjacent bands. This was the output for four (4) specimens. The model correctly selected two (2) of these specimens and incorrectly categorized the remaining two. In this case the errors were conservative, but this is likely to be due to random chance and not inherent conservatism in the model.

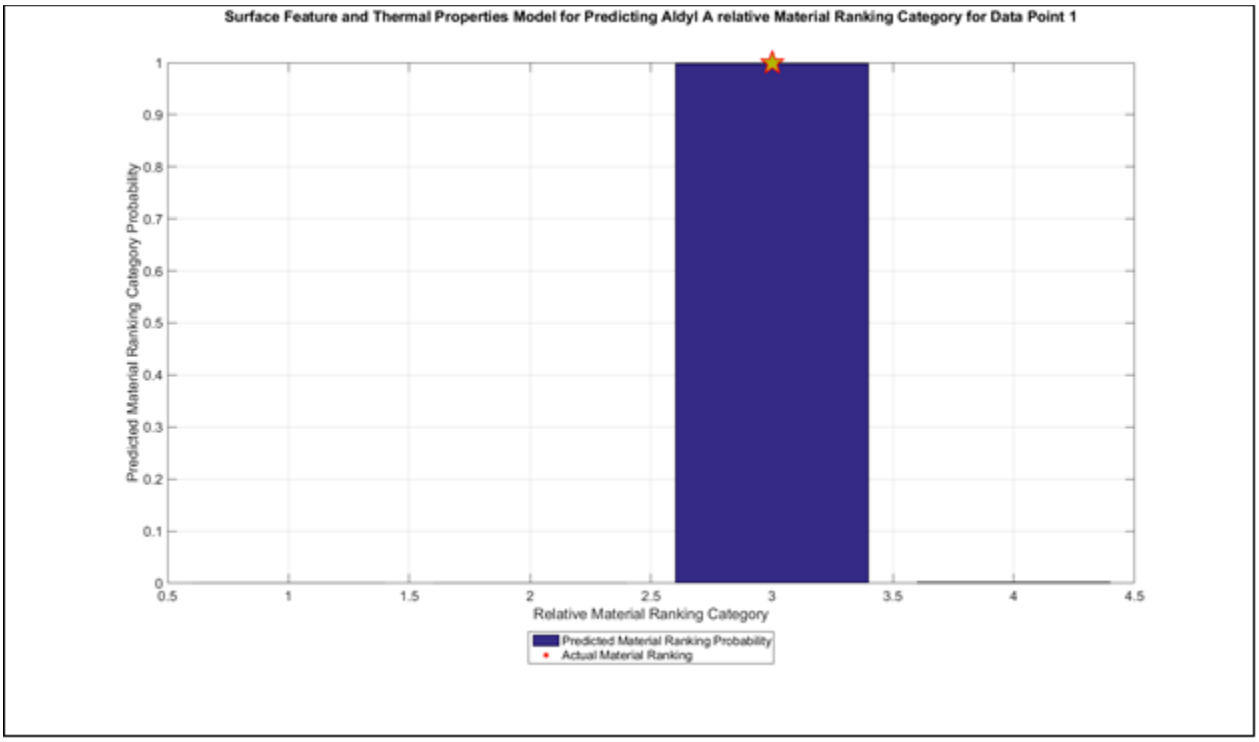


Figure 1-18. Probabilistic prediction of RPM performance band for data point 1

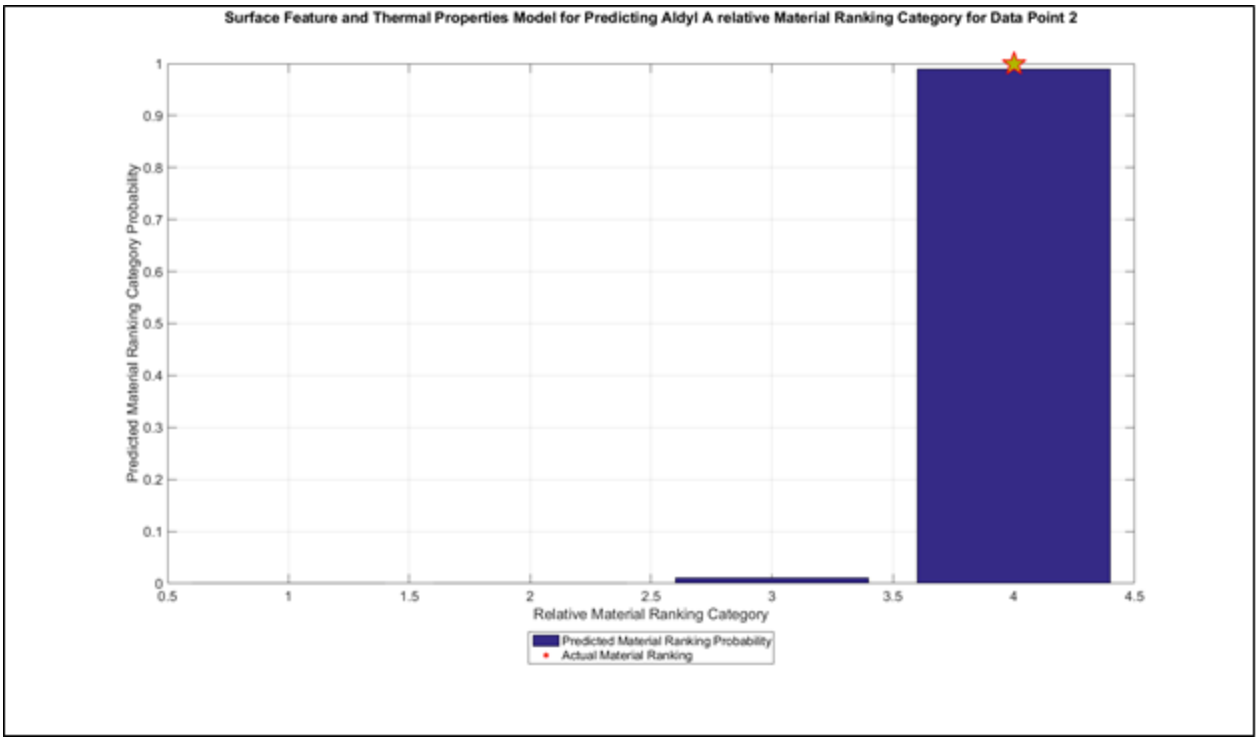


Figure 1-19. Probabilistic prediction of RPM performance band for data point 2

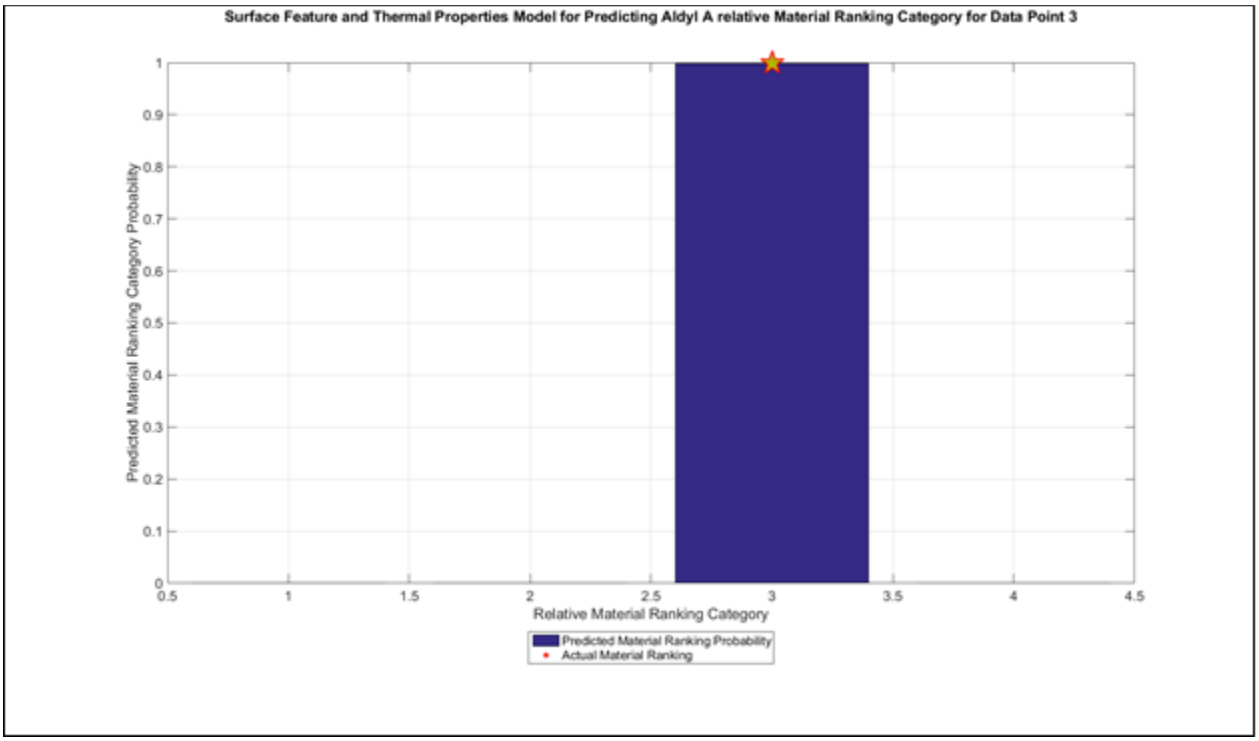


Figure 1-20. Probabilistic prediction of RPM performance band for data point 3

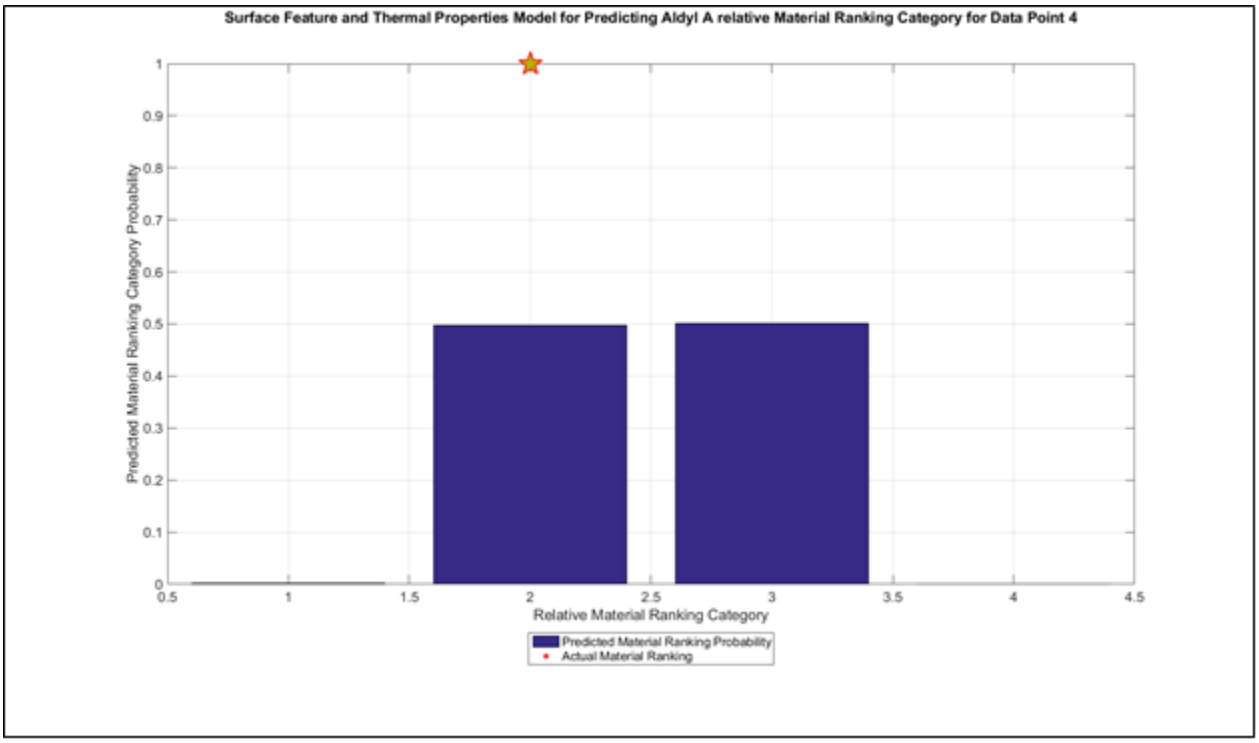


Figure 1-21 Probabilistic prediction of RPM performance band for data point 4

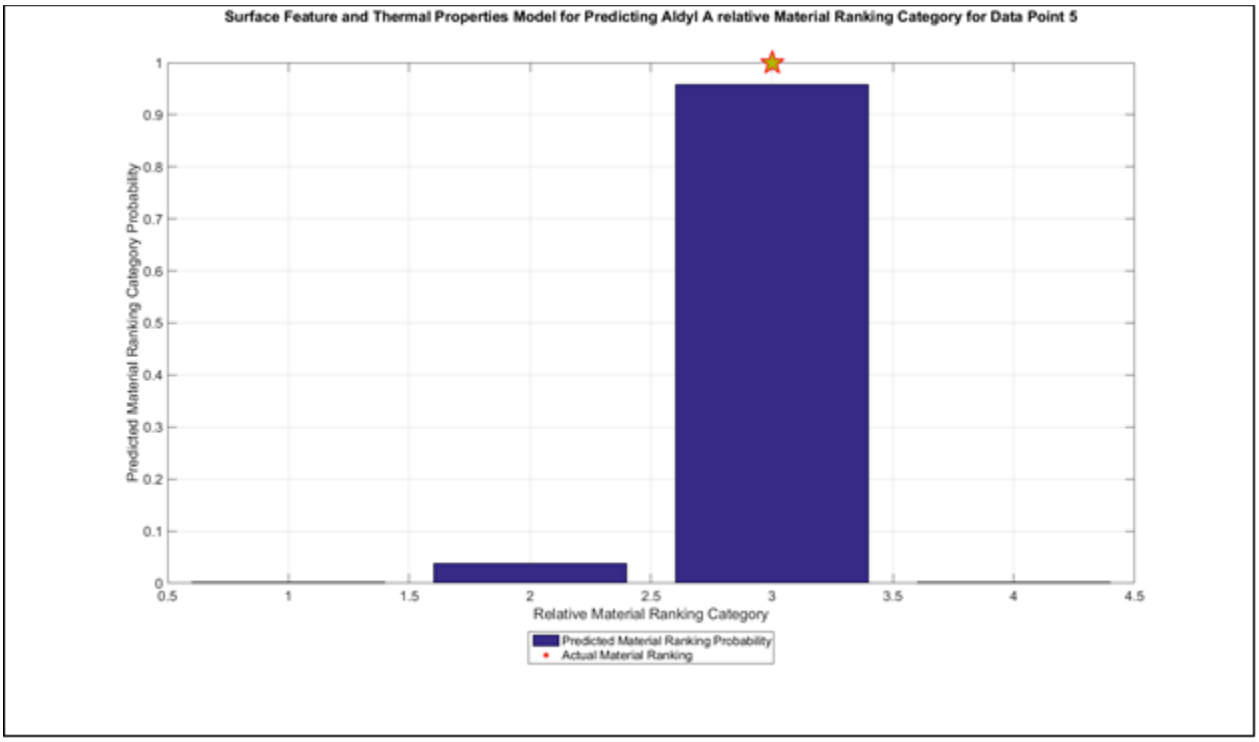


Figure 1-22. Probabilistic prediction of RPM performance band for data point 5

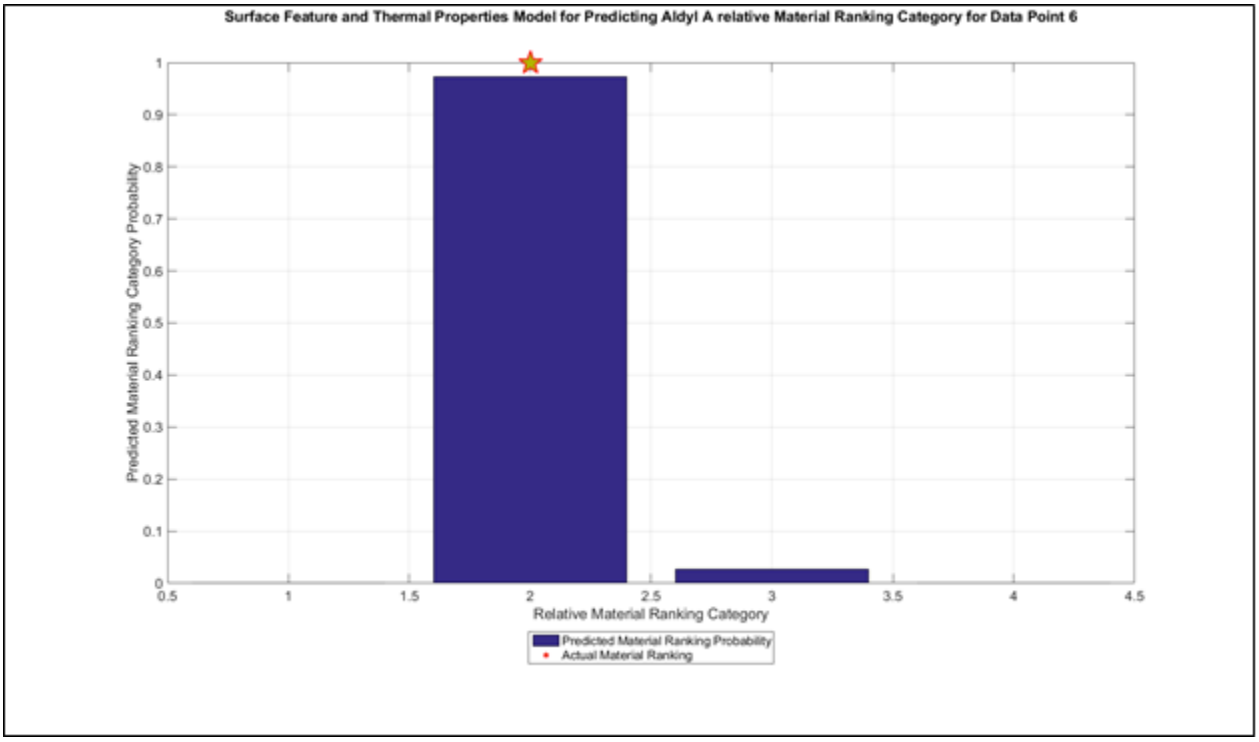


Figure 1-23. Probabilistic prediction of RPM performance band for data point 6

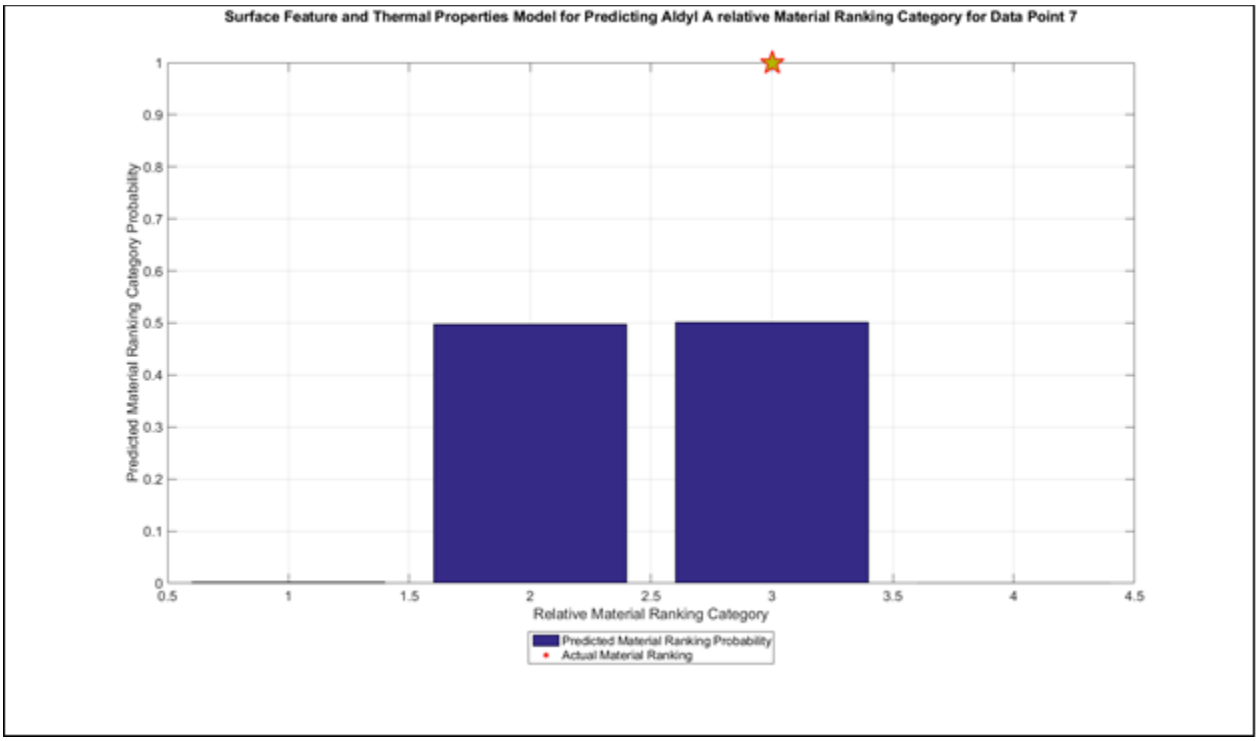


Figure 1-24. Probabilistic prediction of RPM performance band for data point 7

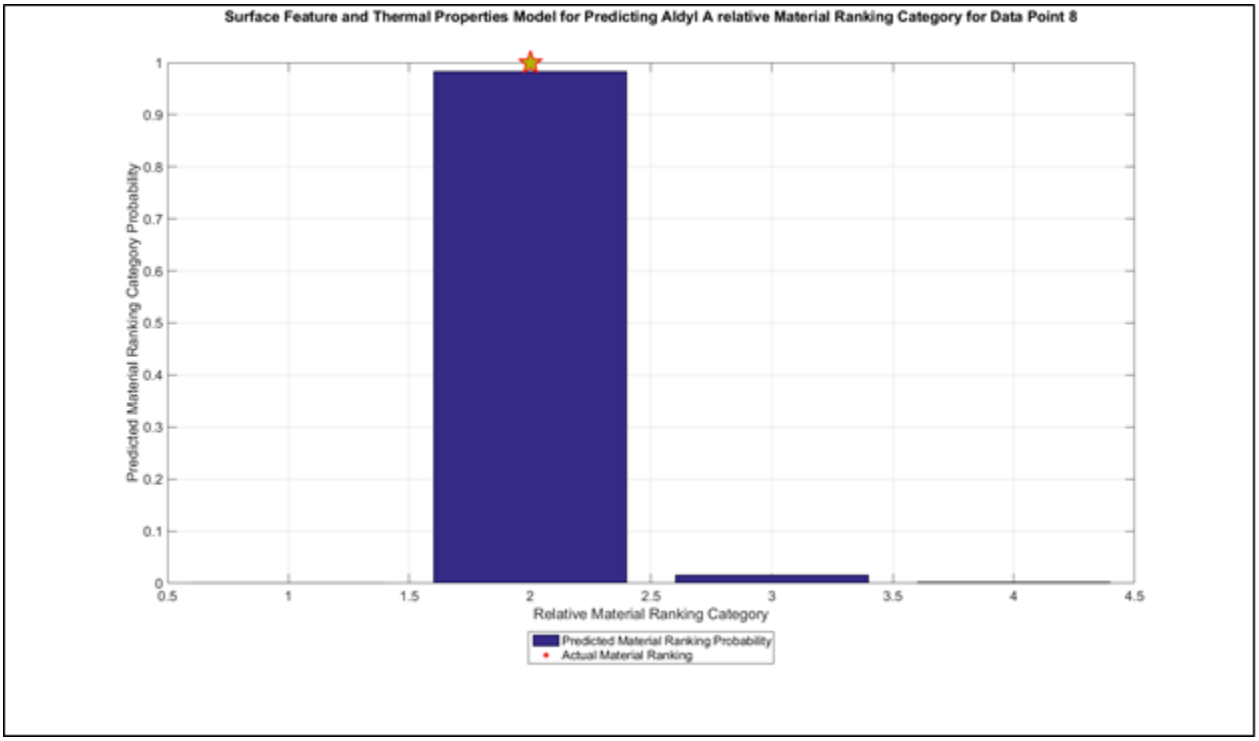


Figure 1-25. Probabilistic prediction of RPM performance band for data point 8

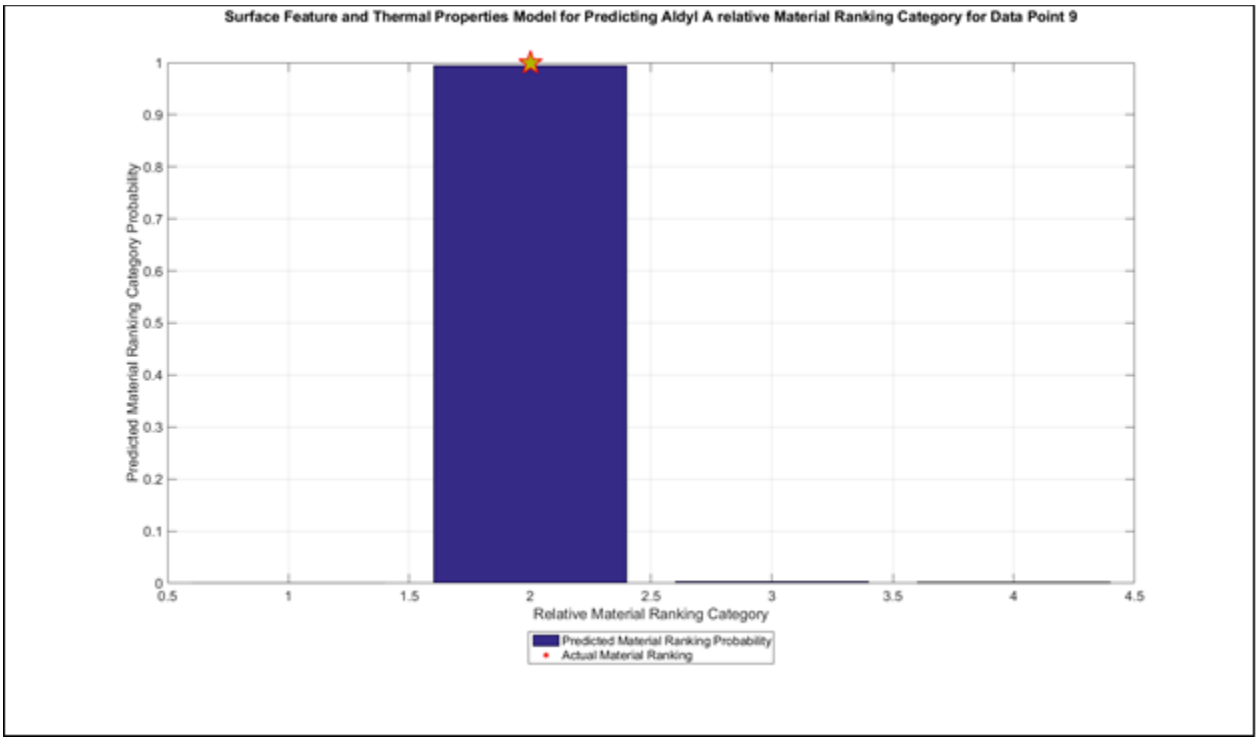


Figure 1-26. Probabilistic prediction of RPM performance band for data point 9

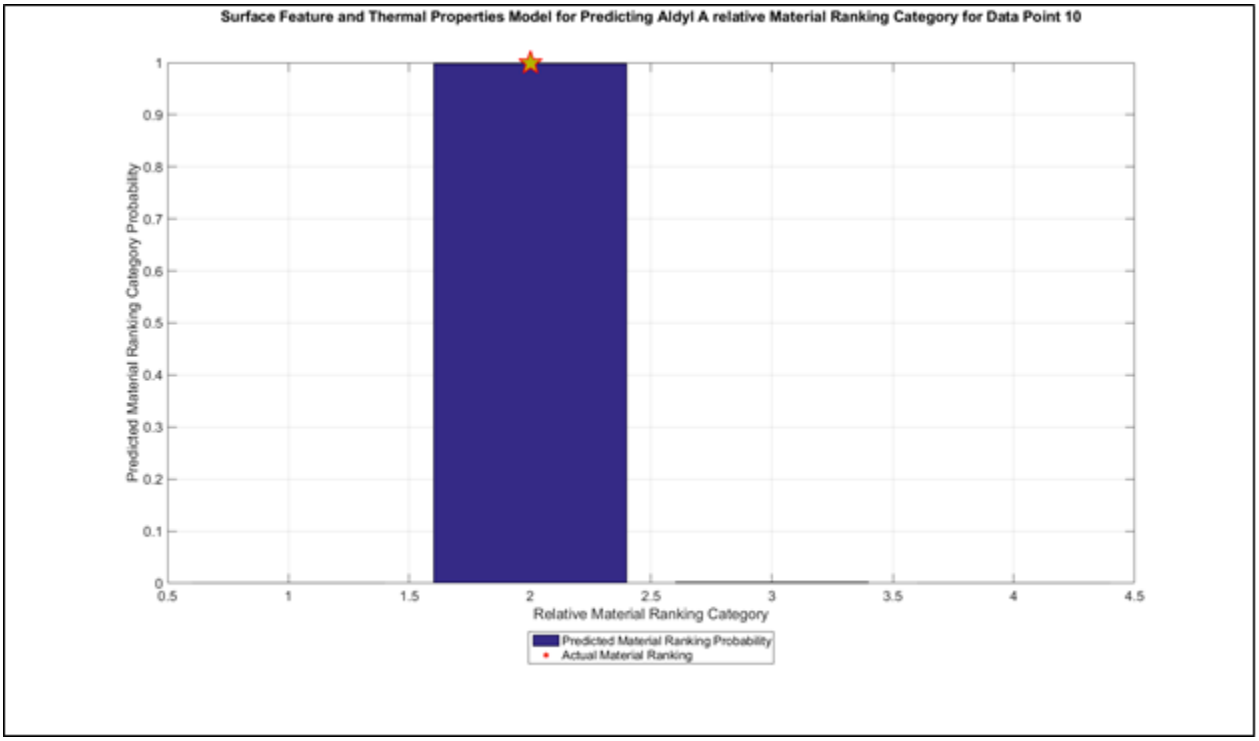


Figure 1-27. Probabilistic prediction of RPM performance band for data point 10

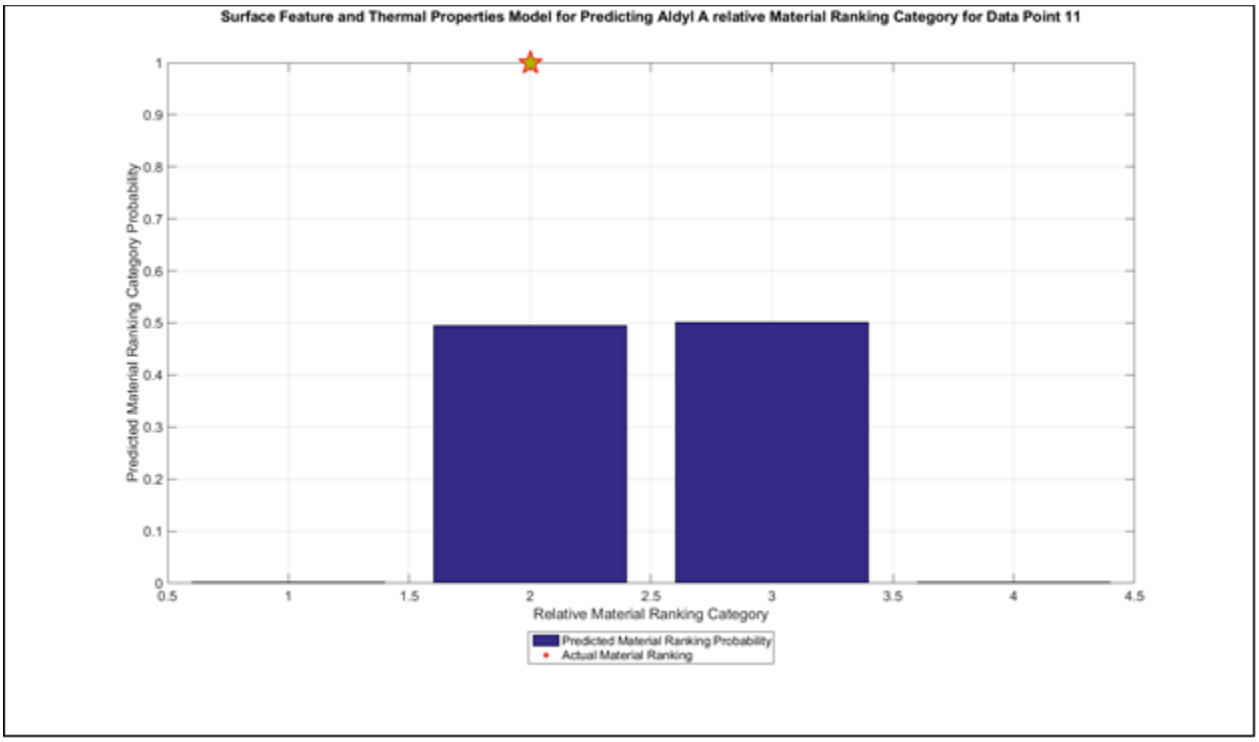


Figure 1-28. Probabilistic prediction of RPM performance band for data point 11

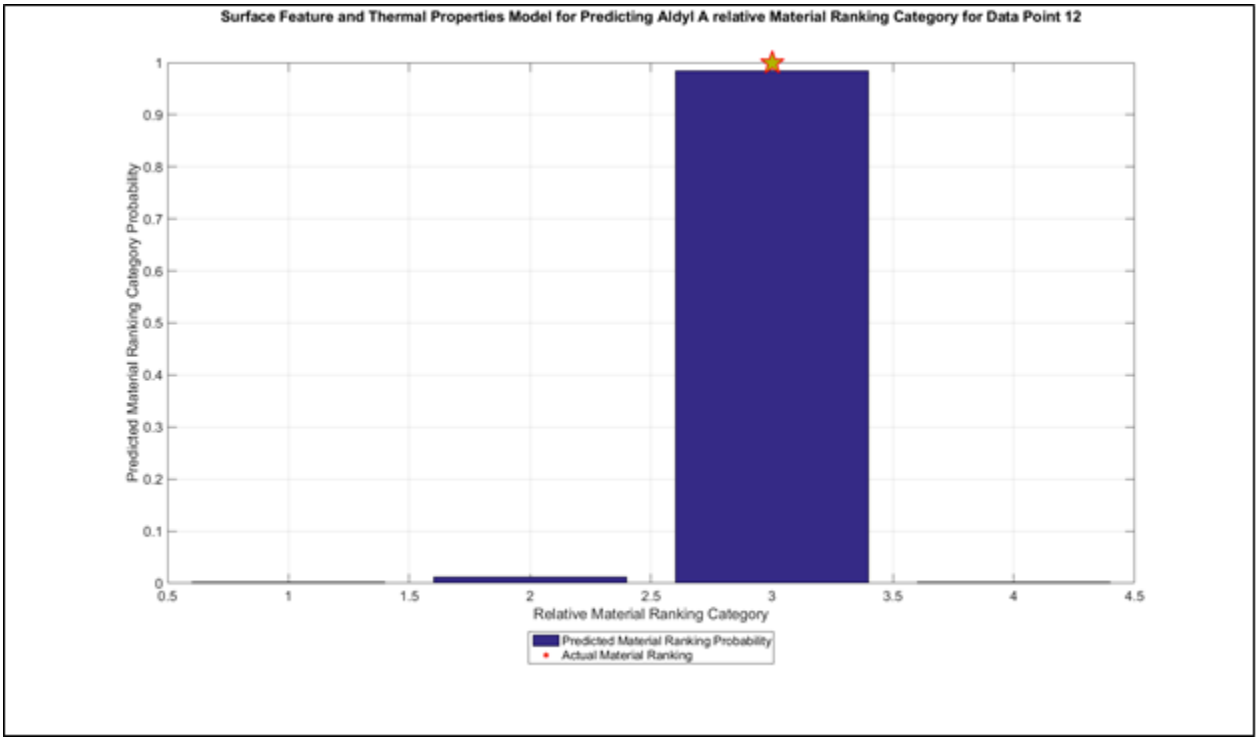


Figure 1-29. Probabilistic prediction of RPM performance band for data point 12

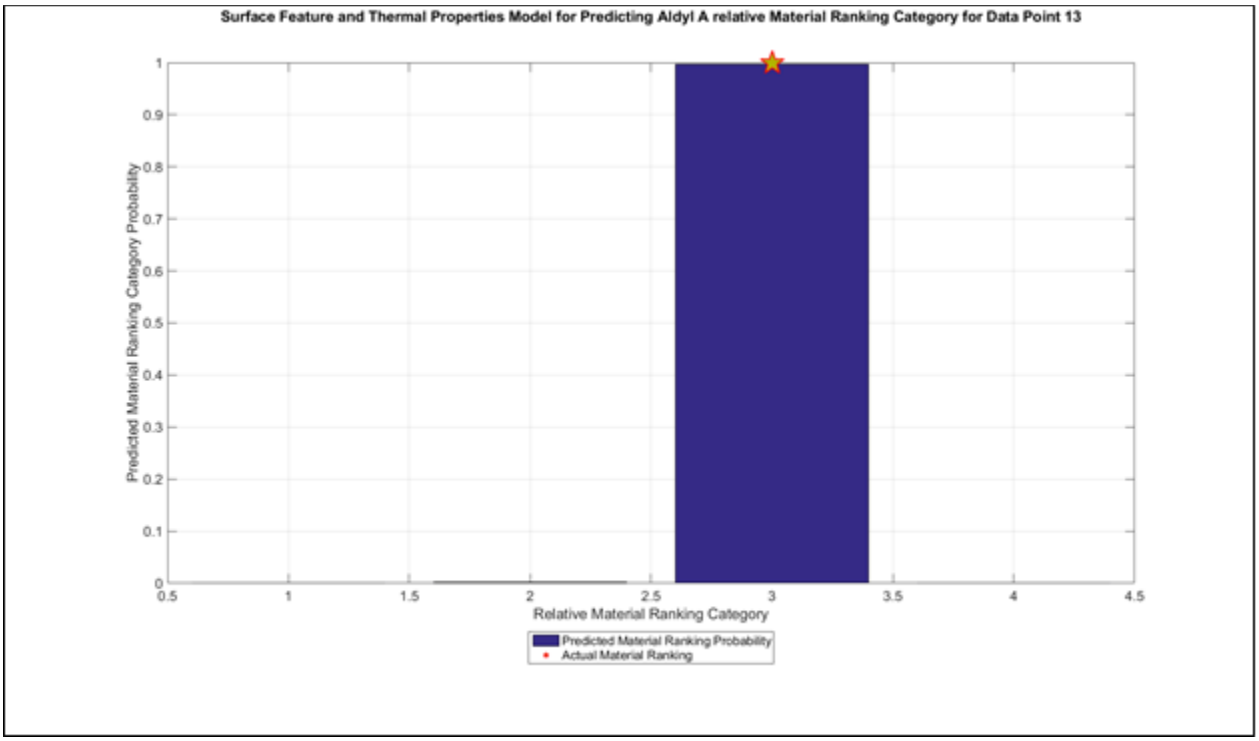


Figure 1-30. Probabilistic prediction of RPM performance band for data point 13

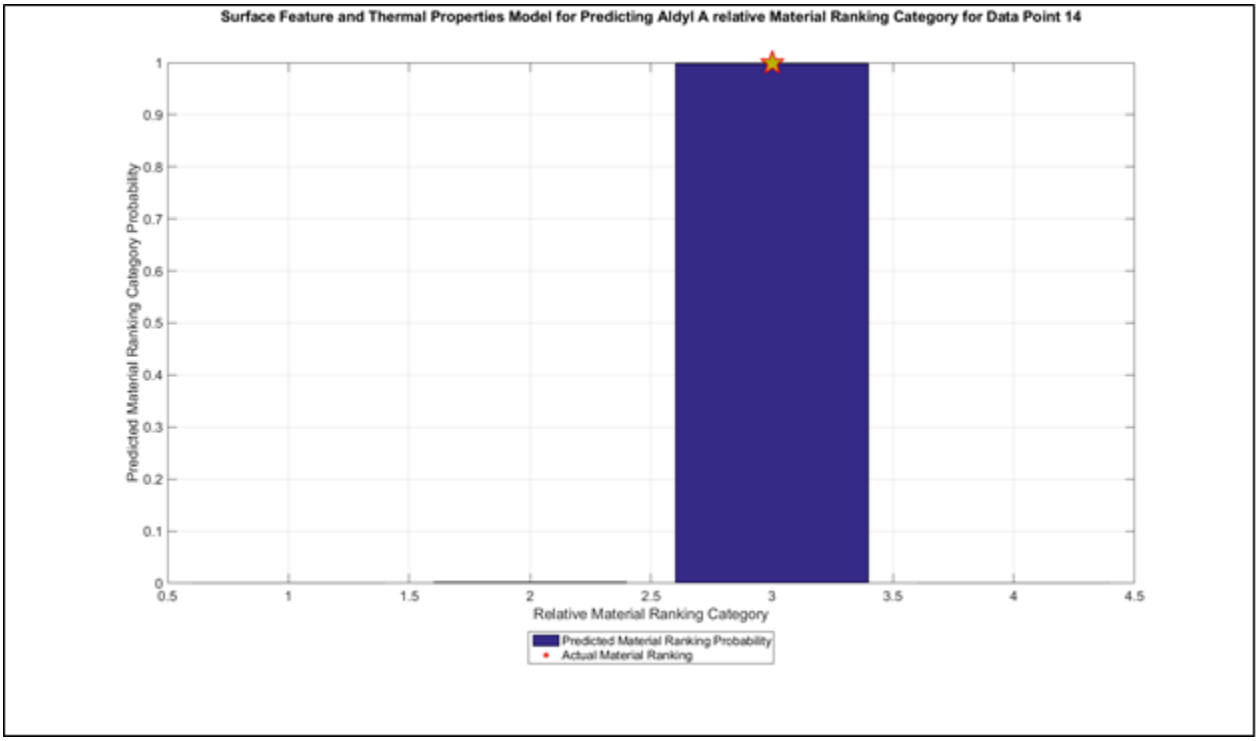


Figure 1-31. Probabilistic prediction of RPM performance band for data point 14

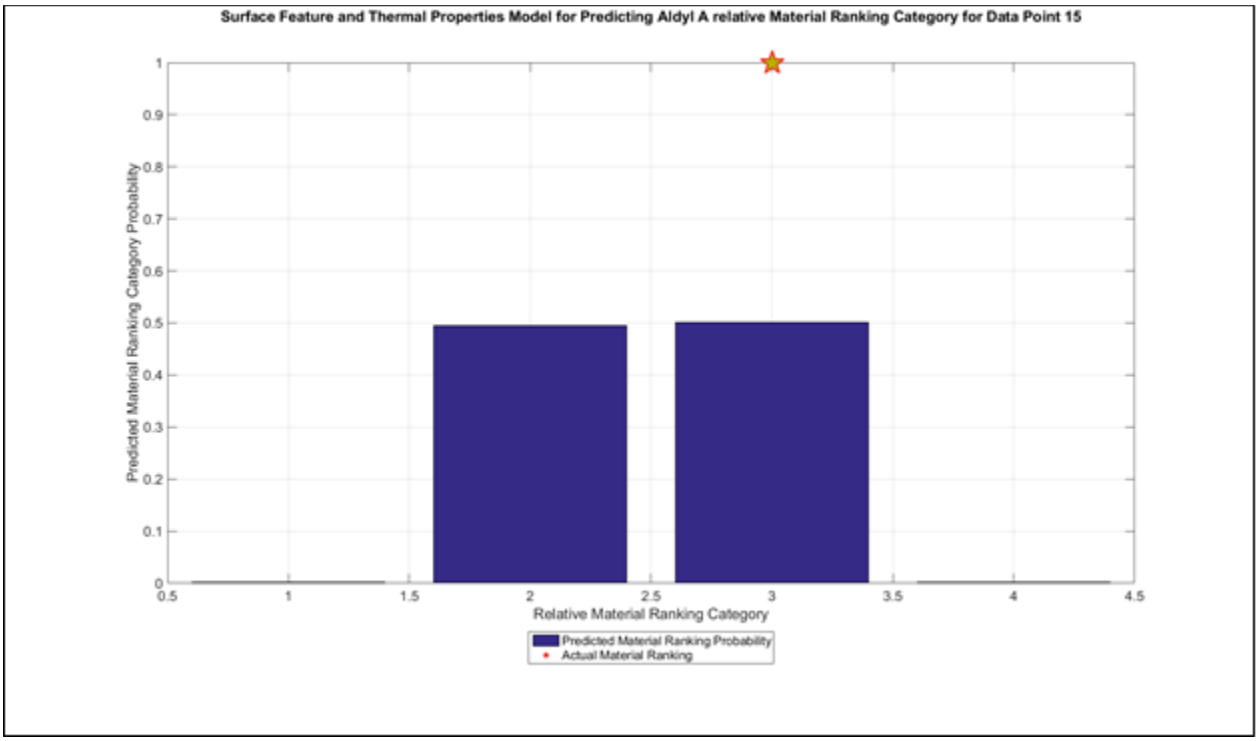


Figure 1-32. Probabilistic prediction of RPM performance band for data point 15

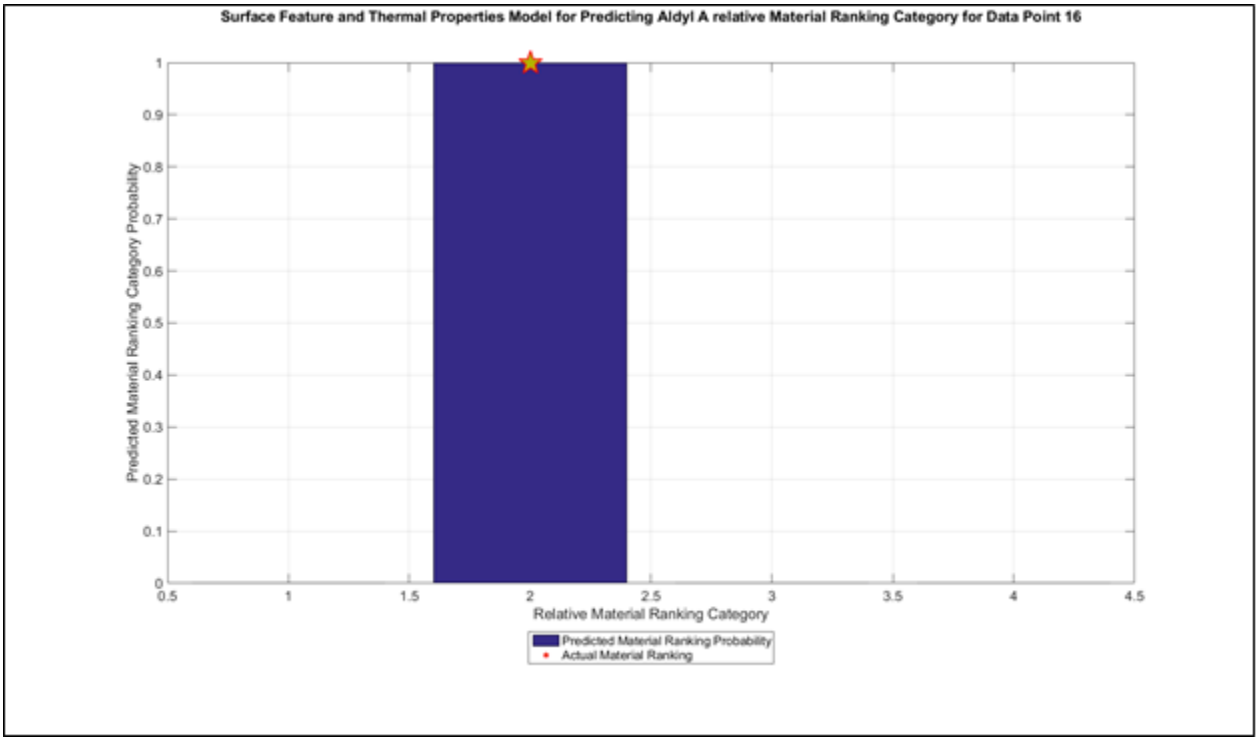


Figure 1-33. Probabilistic prediction of RPM performance band for data point 16

Applying the Surface Correlation Model

The usefulness of the surface correlation model in reducing predictive uncertainty is illustrated in the example presented below. Three data points falling into three different RPM bands as ranked by the long-term testing result were selected. The points are highlighted in **Table 1-5**.

Table 1-5 – Data for Evaluating Use of the Model (Table sorted by stress)

LIMS	Material Ranking	23°C Stress [psi]	23°C Time [y]	Location	Vintage
152246-014	2	401	134	Data Set 2	1970
152244-004	2	410	110	Data Set 4	1971
152243-052	4	469	10	Data Set 2	1970
152243-053	3	471	16	Data Set 2	1970
152246-009	2	507	61	Data Set 2	1970
152244-012	2	511	62	Data Set 4	1971
152299-013	2	514	39	Data Set 1	1969
152245-025	2	514	41	Data Set 3	1974
152299-009	3	556	7	Data Set 1	1969
152299-003	2	556	27	Data Set 1	1969
152299-011	3	605	6	Data Set 1	1969
152244-006	3	620	11	Data Set 4	1971
152299-010	3	827	6	Data Set 1	1969
152244-013	3	963	11	Data Set 4	1971
152299-006	3	989	2	Data Set 1	1969
152243-004	3	1164	2	Data Set 2	1970

The material rankings for the three samples are 2, 3 and 4. The RPM data points for the samples with rankings 2 and 4 fall in the range of test stresses that correspond to pipe operating at 45 psig with stress risers in the 2.2 to 3.1 range (very common in installations).

The reference temperature used in this example is 23°C as it is the standard reference temperature used in many baseline calculations. The same method would be used for any chosen reference temperature (e.g., 15°C, or 10°C). The predicted material rankings are constant across all temperatures – they simply need to be applied to the correct reference temperature graph for the analysis at hand.

The region of the graph being discussed is highlighted with a shaded ellipse in **Figure 1-34.** The RPM test data indicate that the expected lifetimes in this stress range can be anywhere from about 3 years to 70 years.

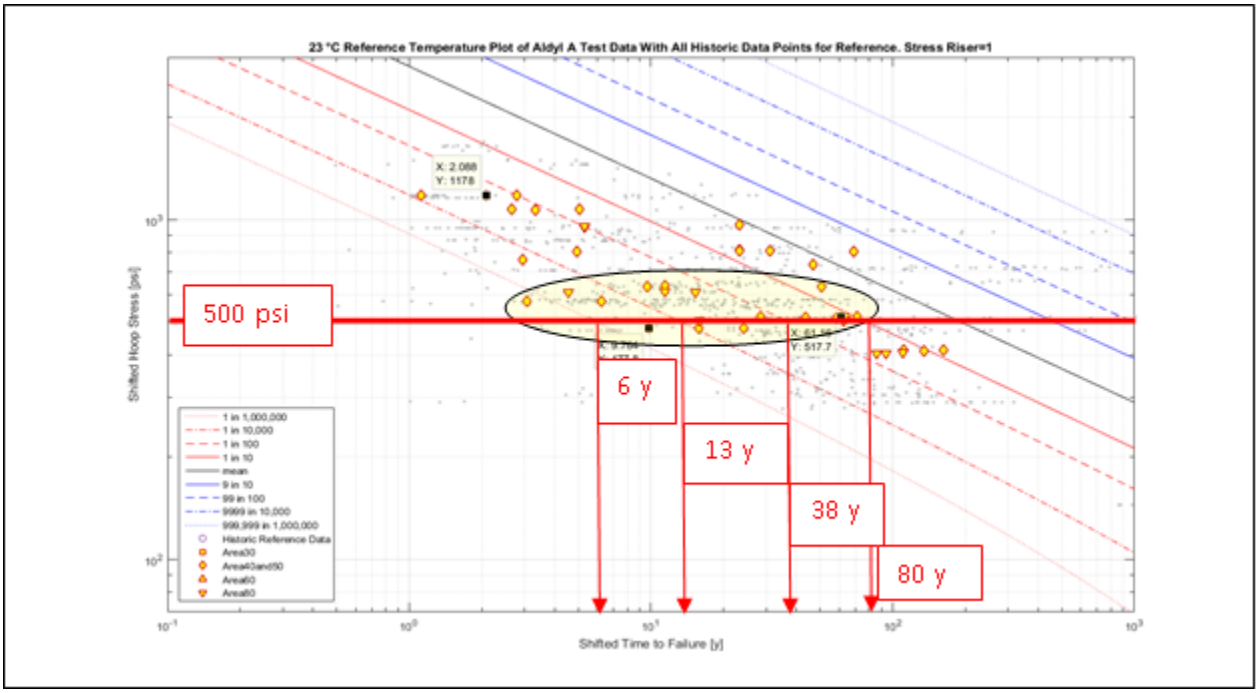


Figure 1-34. Highlighting Correctly-Predicted Reference Categories for Three Example Points

If we take advantage of the material ranking prediction the expected residual lifetime ranges are narrowed as follows for the case of a pipeline operating at 45 psig (225 psi hoop stress for DR 11 pipe) and at a stress riser of 2.2 (=500 psi/225 psi):

1. Material Ranking 2: 38-80 years, 42-year range.
2. Material Ranking 3: 13-38 years, 25-year range.
3. Material Ranking 4: 6-13 years, 7-year range.

The expected lifetime prediction is now more precise and gets more specific as the relative material ranking moves towards a higher likelihood of failure. This is due to the power law relationship between stress and time to failure – when plotted on a log-log set of axes we get straight lines, but each decade is an order of magnitude shorter than the decade to the right along the x-axis and the gradations are not linear.

Surface Correlation Conclusions

1. The revised version of the surface-feature model is a significant improvement over the previous version.
2. The surface feature analysis needs to be expanded to the full RPM data set in future work as well as the retained samples from the OTD 2.8.d (20649) dataset. The combined data sets would give approximately 200 points that can be properly divided into training, check and validation data sets for a "production" model that can be reliably used in system integrity models.

Bi-directional Shift Factors

In the examples above we have presented entire data sets that were generated at multiple temperatures in accordance with the RPM, at single reference temperatures. This is achieved by applying bi-directional shifting to each data point [8-11]. We will now discuss how to develop these bi-directional shift factors, how they relate to the RPM, and creep in polyethylene in general.

The molecular Structure of Polyethylene

We noted above that the model that Coleman [5] selected was that of rate process theory, where the number of jumps away from their nearest neighbors that elements of the polymer chain take, is the process whose rate we are measuring. We will now explore this concept in a little more detail with respect to polyethylene molecular structure. **Figure 1-35** shows the basic structure of a polyethylene molecules that are an approximately linear chain of ethylene monomers.

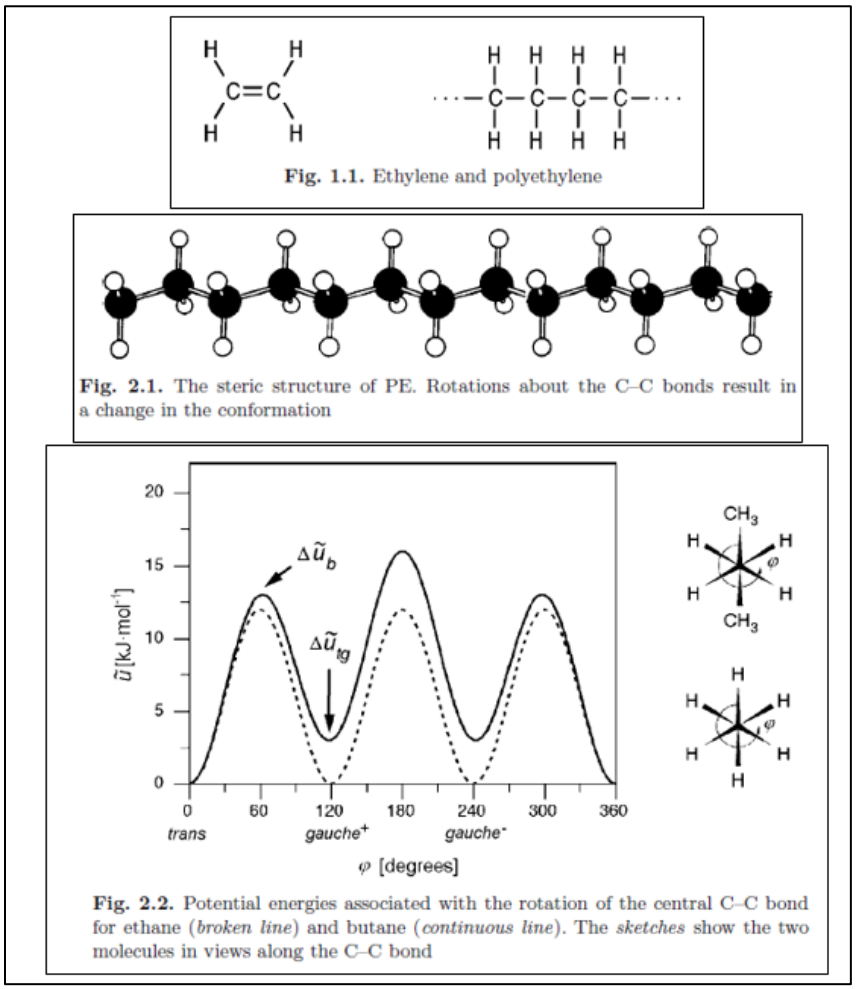


Figure 1-35. Polyethylene Structure and Potential Energies for Rotation of C-C Bond [12]

It is possible for the C-C elements to rotate about their axis. The lower sub-figure in **Figure 1-35** shows the potential energy associated with each of these rotations. These rotations are one of the basic ways in which we can get "local" reorganization of an individual polyethylene molecule. In the solid-state there are many polyethylene chains in close proximity to one another and many more modes for reorganizing the structure come into effect. **Figure 1-36** and **Figure 1-37** show the unit cell structure and a possible mechanism for long chains to order themselves into the semi-crystalline structure typical of medium and high-density polyethylene used in gas distribution piping systems.

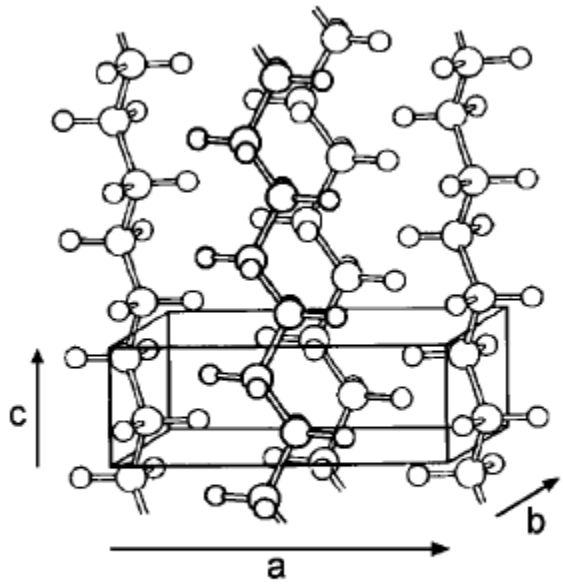


Fig. 5.12. Unit cell of PE crystallites: Orthorhombic symmetry, $a = 7.42$ Å, $b = 4.95$ Å, $c = 2.55$ Å. Each cell is occupied by two C_2H_4 groups. Structure determination by Bunn [39]

Figure 1-36. Unit Cell Structure of PE Crystallites [12]

Fig. 5.7 The 'solidification model' of the crystallisation process, showing how a chain can be incorporated into a lamellar structure without significant change of overall shape. (Reproduced by permission of IUPAC.)

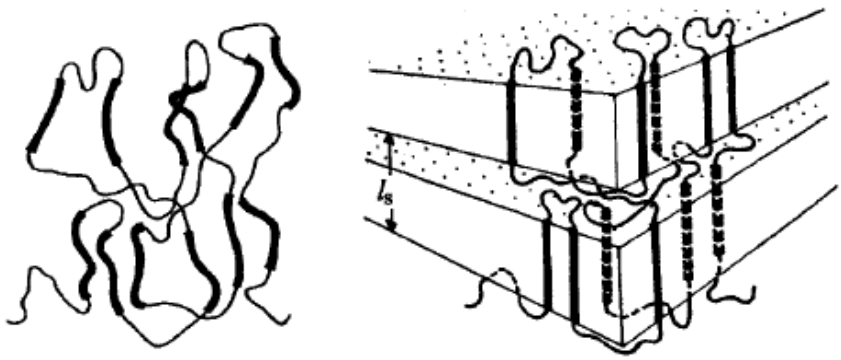


Figure 1-37. Solidification model of Crystallization Process [13]

There are several methods available to us for probing the mechanical and electromagnetic responses of polyethylene under cyclic loading. Nuclear Magnetic Resonance (NMR) and Dynamic Thermo-Mechanical Analysis (DTMA) techniques are very effective in measuring the activation energies for various relaxation phenomena in the solid state. **Figure 1-38** shows some results derived from these methods for polyethylene.

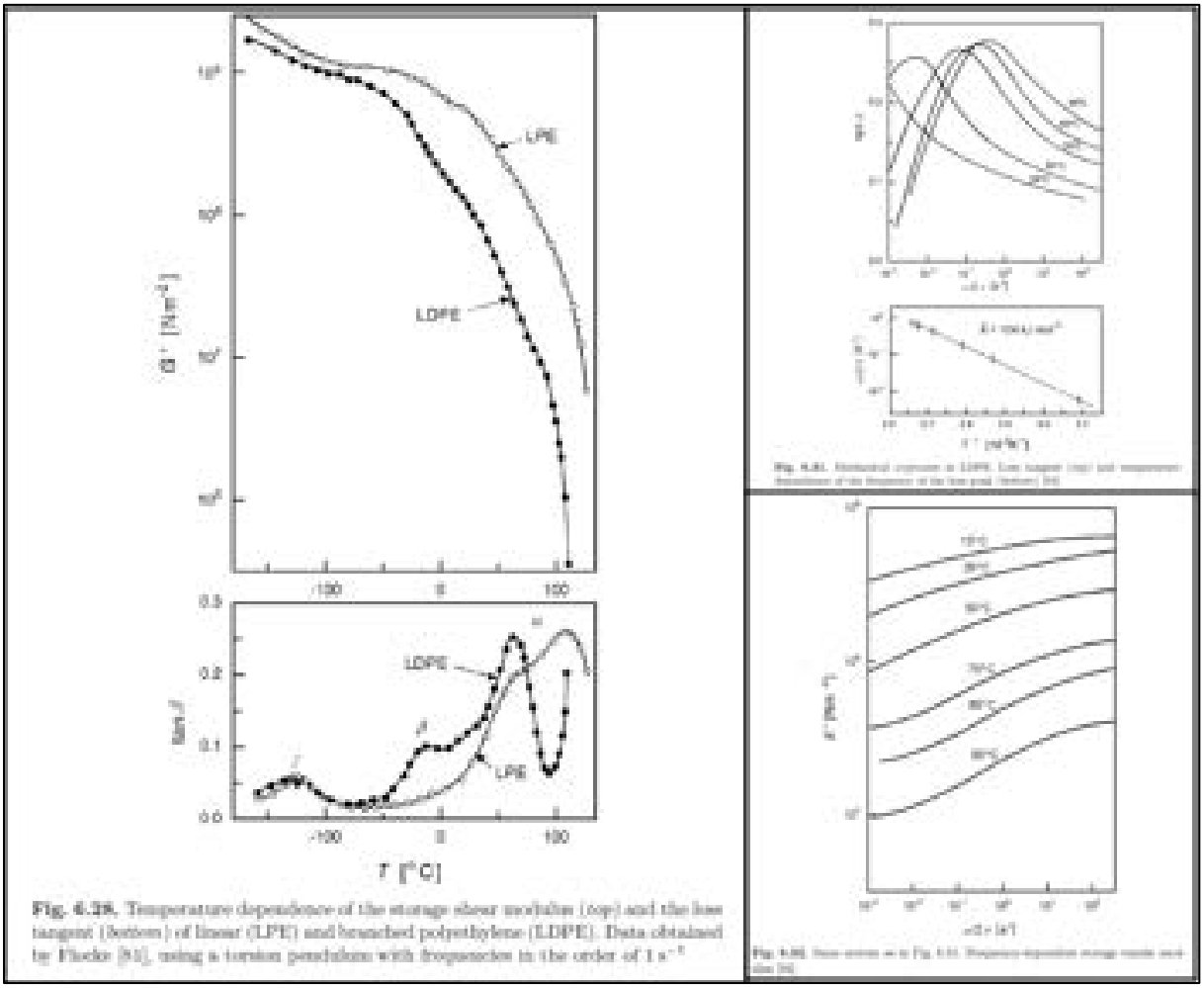


Figure 1-38. Left-Relaxation Mechanisms in PE, Upper Right – Mechanical α-Process Loss Tangent as a Function of Frequency, Lower Right - Mechanical α-Process Storage Modulus as a Function of Temperature [12]

The relaxation of interest to us is the α-process. This is the molecular movement process that Coleman [5] was alluding to. Strobl [12] provides an excellent description of this process and the energies (measured by NMR and DTMA) associated with it in the insert below.

The largest changes in the mechanical properties of polyethylenes with moderate to high crystallinity are caused by the α-process. Figures 6.31 and 6.32 present results of frequency-dependent measurements of the tensile modulus of such a sample, conducted at different temperatures between 26°C and 95°C. As can be seen, the loss tangent shows a systematic shift to higher frequencies. The temperature dependence of the loss maximum is indicative of an activated process with an activation energy A = 104 kJ mol^{-1}. Figure 6.32 shows that the storage tensile modulus decays with both decreasing frequencies and rising temperatures, whereby the latter effect is caused by the continuous melting.
What is the origin of the α-process? Different observations must be included in the considerations. First, remember the results of the NMR experiment presented in Sect. 5.4.2. ***Here, a longitudinal chain transport through the crystallites was clearly indicated. The chain motion is apparently accomplished by a 180°-twist defect, which is created at a crystal surface and then moves through the crystallite to the other side. As a result, all monomers of a crystalline sequence are rotated by 180° and shifted over the length of one CH_2-unit. This screw-motion alone cannot set up the α-process, since it is mechanically inactive. As the crystals remain unchanged, both internally and in their external shape, there is no coupling to a stress field. Furthermore, the high relaxation strength of the α-process suggests a location in the weak amorphous parts of the structure rather than in the crystallites. How can the different observations be cast in one common picture? The answer is that the α-process in polyethylene has a composite nature. The mechanical relaxation indeed originates from an additional shearing of the amorphous regions. However, the prerequisite for this shearing is a chain movement through the crystallites. If such a motion is thermally activated, the pinning of the amorphous sequences onto the crystallite surfaces is no longer permanent. This allows a reorganization of the amorphous regions, which gives rise to a further stress decay. Hence, in the α-process, two relaxation processes, one in the crystallites and the other one located in the amorphous zones, are coupled. As indicated by the broadness of the loss curves, the α-process is based on a larger group of relaxatory modes. Since the temperature variation leaves the shape of the loss curves essentially unchanged, we conclude that all modes employ the same elementary process, to be identified with the step-like longitudinal shifts of the crystalline sequences. NMR experiments and dynamic mechanical measurements indeed yield nearly identical activation energies – A = 105 $kJmol^{-1}$ in Fig. 5.53 and A = 104 $kJmol^{-1}$ in Fig. 6.31. Of interest is a comparison of the rate of elementary steps with the mechanical relaxation rate. The difference amounts to four orders of magnitude, telling us that the reorganization of the chains in the amorphous regions is a complex procedure requiring a huge number of elementary steps.***

Strobl, G.R., The Physics of Polymers: Concepts for Understanding Their Structures and Behavior. 2013: Springer Berlin Heidelberg. ***pp280-283***
Emphasis added

Bower [13] has a concise explanation of the site-model theory and a derivation of the rate process formulation for relaxation times described by the Arrhenius equation[1]:

$$k = A\, e^{\frac{-E_a}{R\,T}} \qquad \textbf{Equation 1-3}$$

k -	The rate constant
T -	The absolute temperature (in kelvins)
A -	Pre-exponential factor, a constant for each chemical reaction that defines the rate due to frequency of collisions in the correct orientation
E_a -	Activation energy for the reaction (in the same units as R T)
R -	Universal gas constant

In the present context, we are not discussing chemical reactions, but rather the coherent movement of groups of molecular units in a semi-crystalline polymer structure like polyethylene. Site-model theory, in its simplest form, describes two sites, each representing a particular local conformational state of the molecule, separated by an energy barrier as shown in **Figure 1-39**.

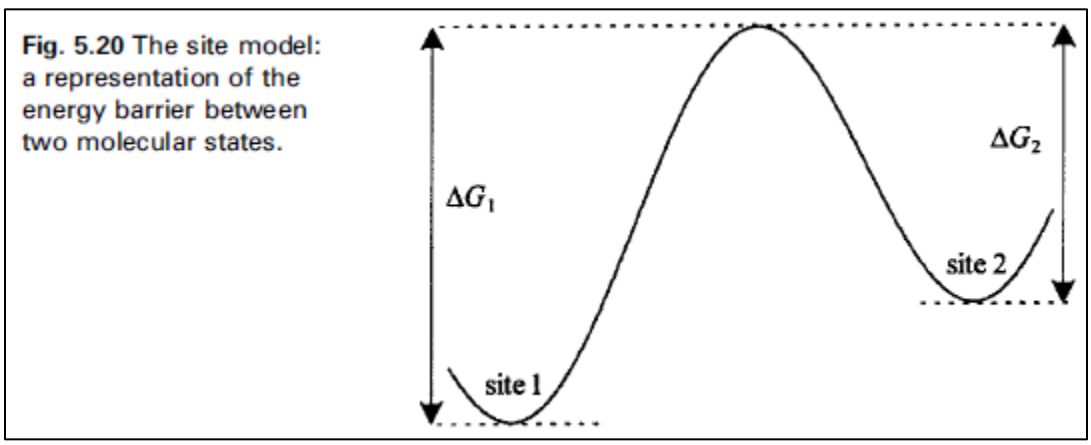


Figure 1-39. Site-model as described by Bower [13] pp 148

Bower discusses the number of molecules in each conformational state and the probabilities that they will transition to the alternate state in a time interval. The endpoint of the derivation is that when a semi-crystalline polymer is subjected to a fixed displacement, the

1 Wikipedia contributors, "Arrhenius equation," *Wikipedia, The Free Encyclopedia,* https://en.wikipedia.org/w/index.php?title=Arrhenius_equation&oldid=799107578 (accessed September 15, 2017
The Arrhenius equation is a formula for the temperature dependence of reaction rates

strain will relax to its equilibrium value with a time constant, or relaxation time, "τ"that is temperature dependent and is equal to the average time between jumps over the barrier. The time constant equation is:

$$\tau = A \exp[\frac{\Delta E}{R\,T}] \qquad \textbf{Equation 1-4}$$

ΔE is the activation energy, which is equal to the enthalpy difference between the two states.

Bower presents an example where he compares the relaxation times at two different temperatures and develops the following equation:

$$\frac{\tau}{\tau_0} = \exp[\frac{\Delta E}{R}(\frac{1}{T} - \frac{1}{T_0})] \qquad \textbf{Equation 1-5}$$

Equation 1-5 is extremely useful as it defines the time-temperature equivalence for a material process that is adequately described by an Arrhenius relationship. The ratio τ/τ_0 is often called a shift factor as it defines a linear multiplier on a known relaxation time at a reference temperature to arrive at a relaxation time at a different temperature:

$$\tau = \tau_0 * \exp[\frac{\Delta E}{R}(\frac{1}{T} - \frac{1}{T_0})] = \tau_0 * a_T \qquad \textbf{Equation 1-6}$$

a_T – Shift factor for time/temperature equivalence

Popelar [10, 11, 14] arrived at horizontal and vertical shift factors for polyethylene empirically and published "universal" shift factors that adequately described the time/temperature equivalence of several commercially available polyethylene materials used in natural gas distribution systems at the time he published his work. These shift factors became known in the industry as the Popelar Shift Factors and their application is ubiquitous. The methodology[2] used in the United States for determining the strength ratings of polyethylene resins uses the Popelar shift factors in the validation of regression curves developed for specific application temperatures.
Mavridis [9] addresses the temperature dependence of polyolefin melt rheology and describes a DTMA methodology for extracting the activation energies needed to calculate temperature dependent shift factors. He discusses thermorheological simplicity of melts, where all

2 http://plasticpipe.org/pdf/tr-3-2017a.pdf
TR-3/2017a HDB/HDS/PDB/ SDB/MRS/CRS Policies: Policies and Procedures for Developing Hydrostatic Design Basis (HDB), Hydrostatic Design Stresses (HDS), Pressure Design Basis (PDB), Strength Design Basis (SDB), Minimum Required Strength (MRS) Ratings, and Categorized Required Strength (CRS) for Thermoplastic Piping Materials or Pipe, **Section F.4.1.2**

relaxation processes have the same temperature dependency. He points out that by independently evaluating the temperature dependence of relaxation times and modulus we can derive coherent vertical and horizontal shift factors that fully describe the relaxation behavior of polyolefins as a function of temperature. The method utilized by Mavridis is simply fitting of DTMA data to **Equation 1-5** to extract ΔE for the horizontal and vertical shift factors.

Lever [8] applied the DTMA method described by Mavridis to multiple pipe materials and showed that there is variation in the shift factors amongst materials as well as batch to batch variation. At this point it needs to pointed out that there is a difference in the Mavridis and Popelar approaches in that Popelar [11] used stress relaxation data to fit a non-linear viscoelastic model that captures the essence of the materials relaxation behavior. This viscoelastic model was used to develop horizontal and vertical shift factors that produced a coherent master relaxation curve for a wide array of polyethylene materials in use at that time. The Popelar shift factors are presented in the form:

$$\text{Shift Factor} = \exp(\text{Constant} * (\mathrm{T} - \mathrm{T_{ref}}) \qquad \textbf{Equation 1-7}$$

T – Test Temperature [°C]
T_{ref} – Reference Temperature [°C]

By contrast the Mavridis methodology fits an Arrhenius form yielding:

$$\text{Shift Factor} = \exp(\frac{\text{Activation Energy}}{R} * (\frac{1}{T} - \frac{1}{\mathrm{T_{ref}}}) \qquad \textbf{Equation 1-8}$$

R – Boltzmann's Constant
T - Test Temperature [K]
T_{ref} – Reference Temperature [K]

Equation 1-7 is dependent on the temperature differential between the test temperature and the reference temperature and is very nearly linear on a log/linear plot. **Equation 1-8** is dependent on the difference between the inverse of the test temperature and the inverse of the reference temperature in degrees Kelvin. This relationship has noticeable curvature on a log/linear plot **Figure 1-40**.

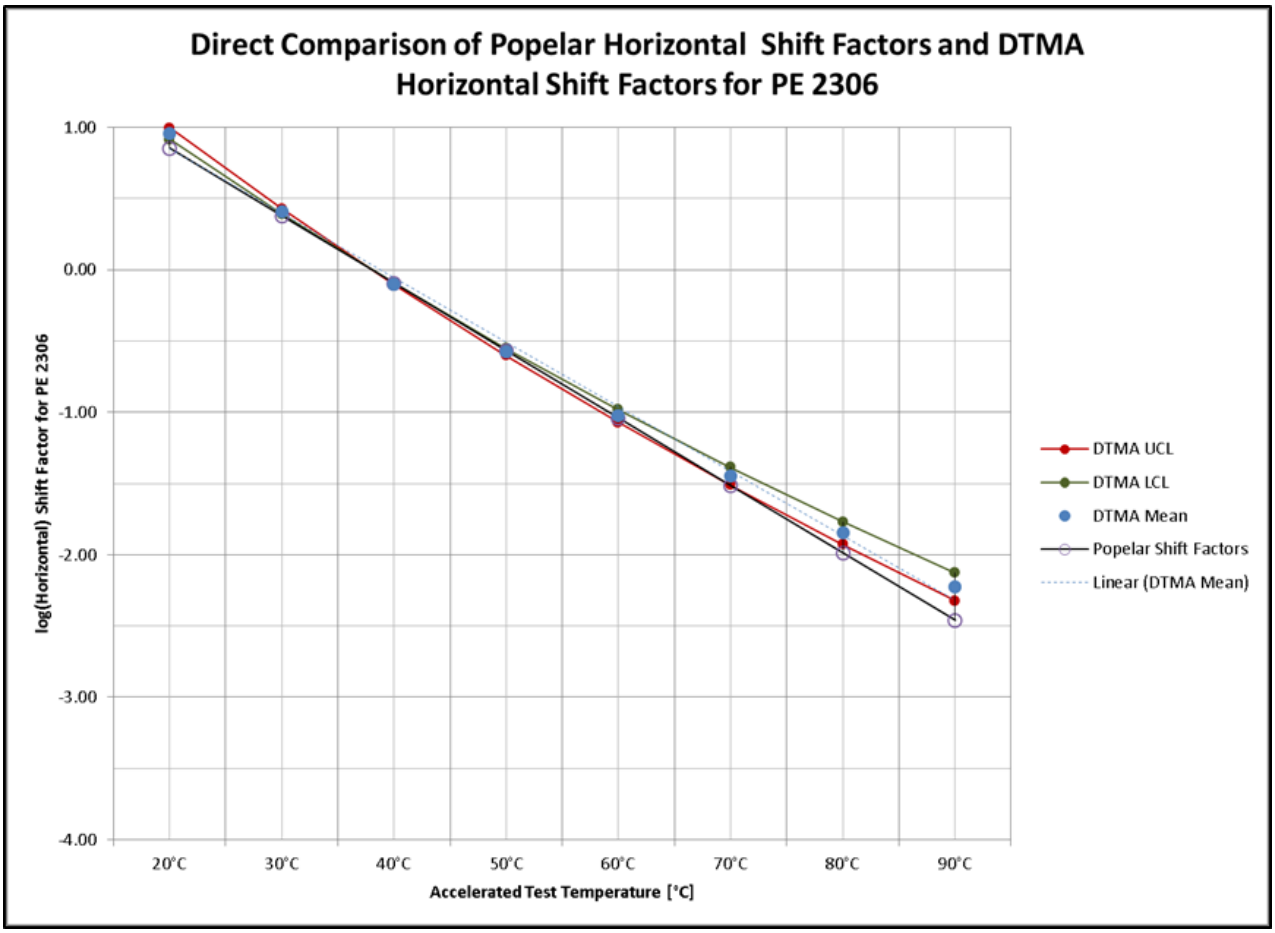
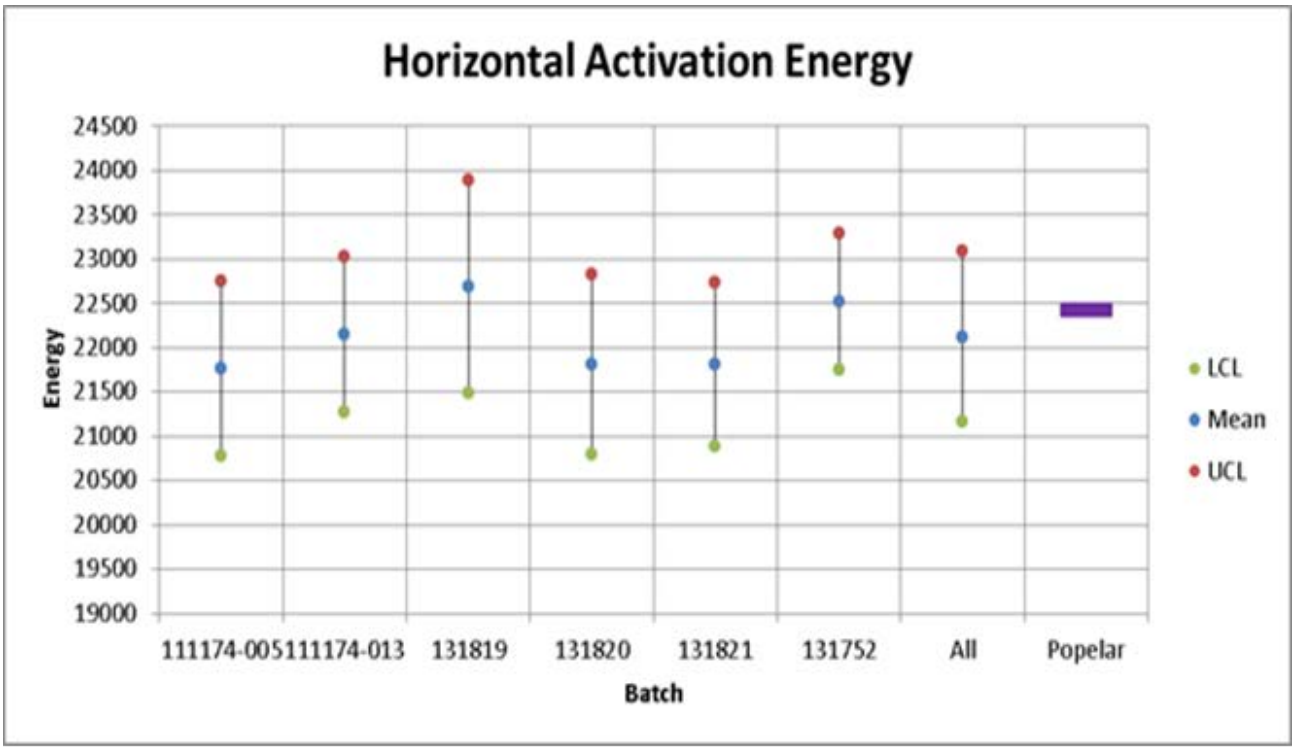


Figure 1-40. Comparison of DTMA and Popelar Horizontal Shift Factors in Form Presented in [11]

The results of the DTMA based activation energy determinations are presented in **Figure 1-41** to **Figure 1-44** and **Table 1-6**. The activation energy for the α relaxation process in polyethylene gas distribution pipe compares favorably to the value of 104 kJ/mol Strobl [12] lists for Low Density Polyethylene (LDPE).



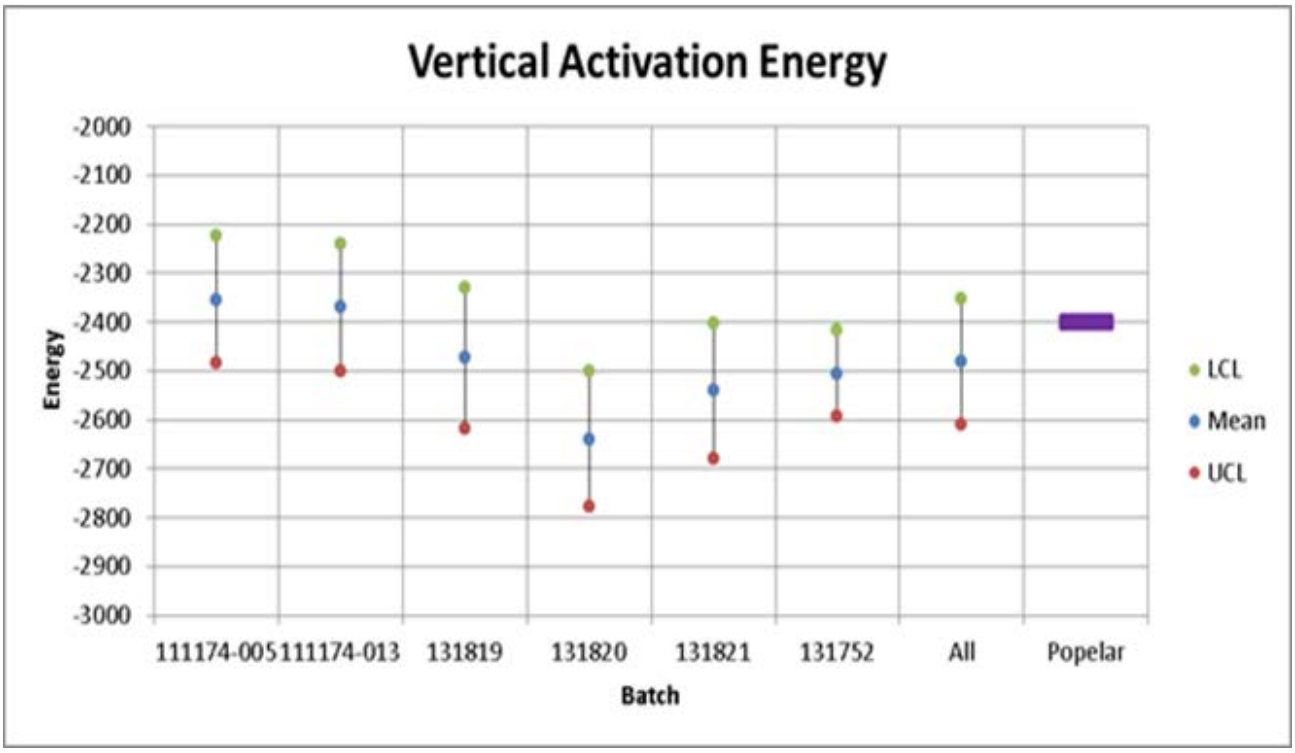


Figure 1-41. Horizontal and Vertical Activation Energies of Aldyl A (PE2306) Pipe Batches Measured by DTMA. Energy in cal/mol

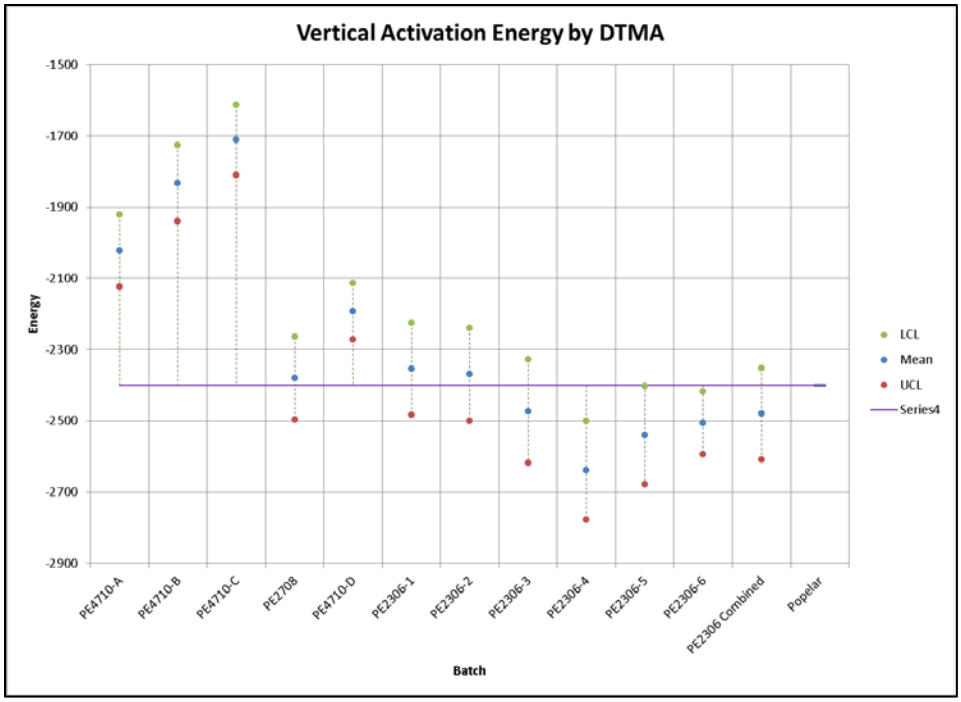


Figure 1-42.Horizontal Activation Energies for Modern and Aldyl A Materials Relative to Equivalent Popelar Activation Energy. Energy in cal/mol

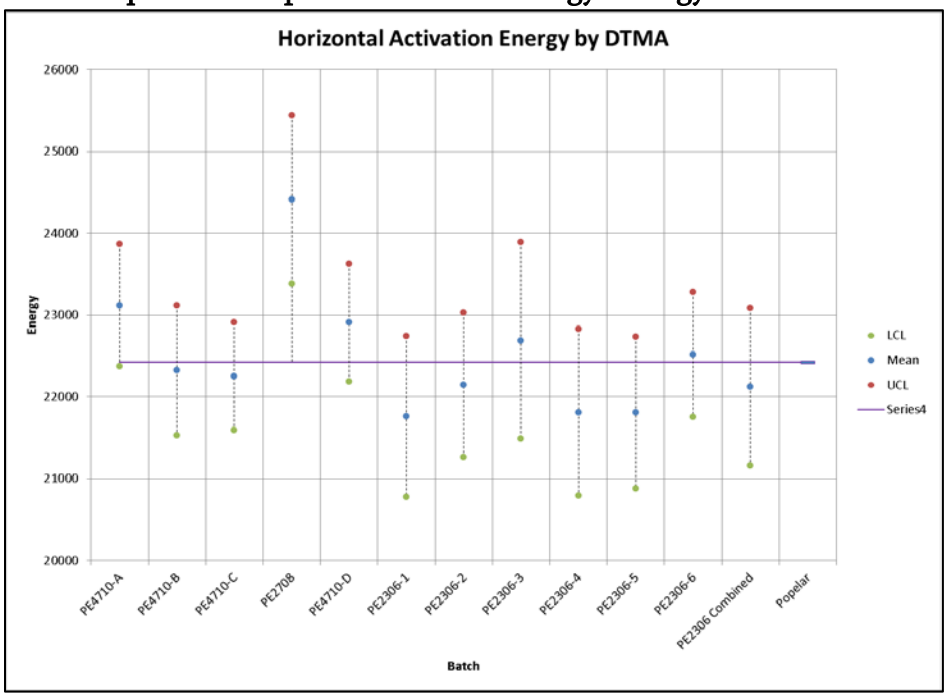


Figure 1-43Vertical Activation Energies for Modern and Aldyl A Materials Relative to Equivalent Popelar Activation Energy. Energy in cal/mol

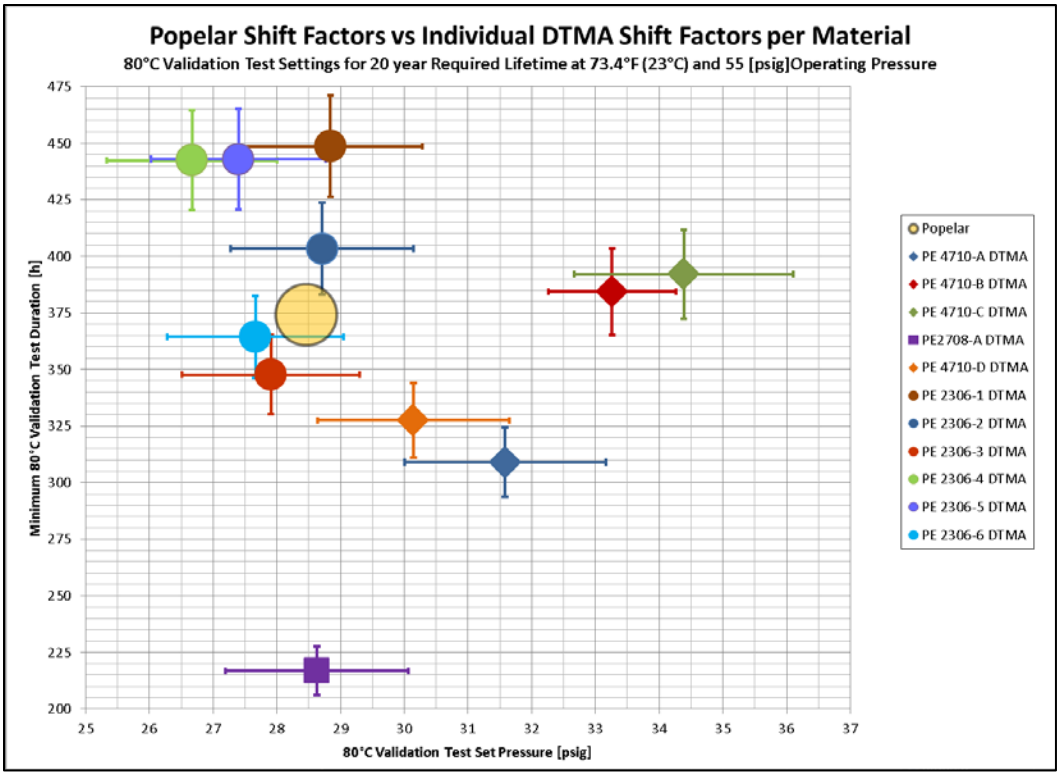


Figure 1-44. 80°C Validation Test time/pressure Combinations for Several Materials to Validate a 20-year Minimum Lifetime at 23°C and an Operating Pressure of 55 psig

Table 1-6. Activation Energies for α relaxation process in polyethylene gas distribution pipe as measured by DTMA

Aldyl A Batch	Total Activation Energy [kJ/mol]		
	UCL	Mean	LCL
111174-005	96	92	87
111174-013	97	93	89
131819	101	95	90
131820	96	92	88
131821	96	92	88
131752	98	95	92
Modern Material	**Total Activation Energy [kJ/mol]**		
	UCL	**Mean**	**LCL**
unimodal MDPE A	114	105	96
bimodal HDPE C	102	98	95
unimodal MDPE B	100	94	89
bimodal MDPE D	104	98	93

Using material specific bi-directional shift factors developed from DTMA measurements allows us to shift data points from tests performed at multiple temperatures to any given reference temperature as shown in **Figure 1-45**. The points in the figure are actual data and the lines reflect the Lower Prediction Limit, Mean and Upper Prediction Limit of the RPM model for the data set. We can see that RPM model and bi-directional shift factors measured for the material tested are coherent. The combination of the two methods is a very powerful data analysis tool.

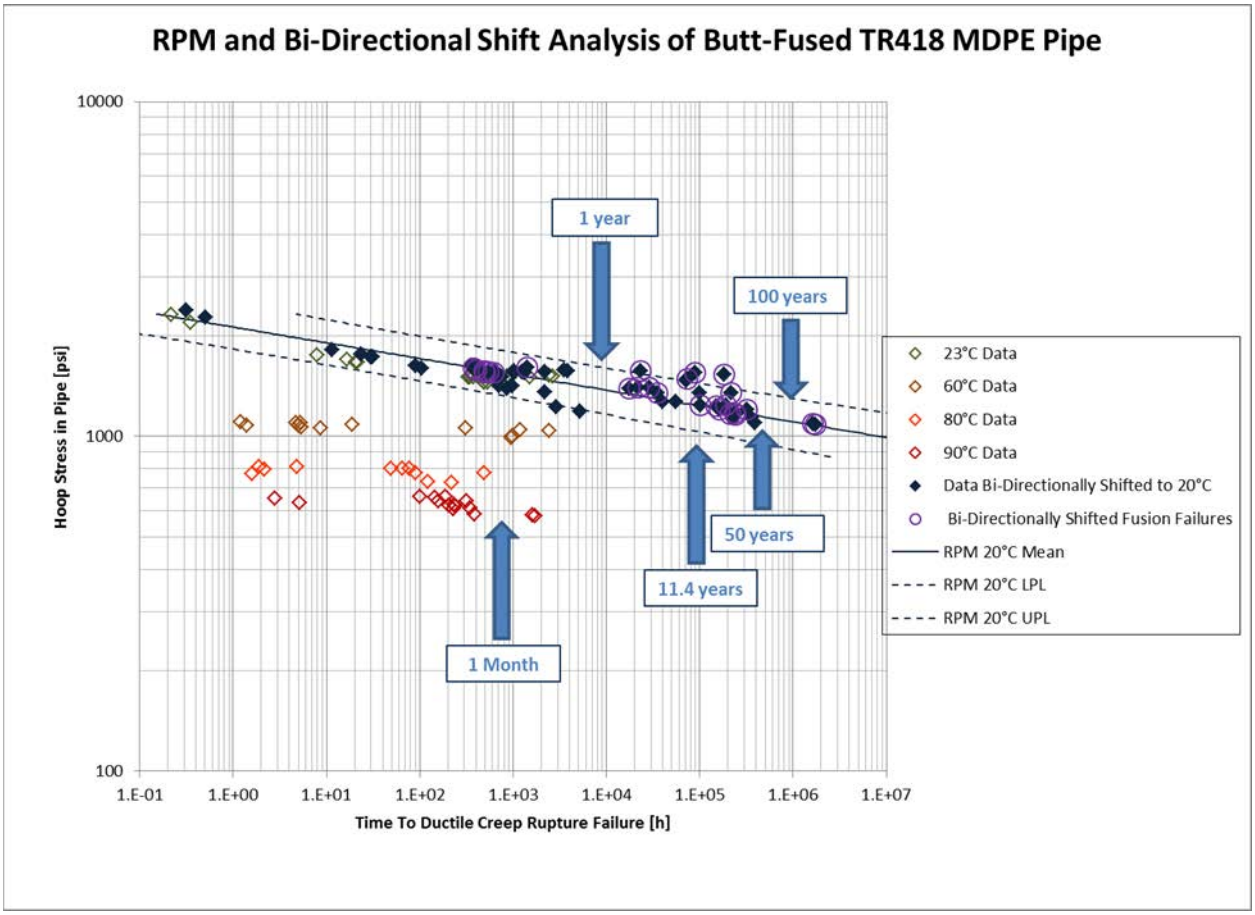


Figure 1-45. Plot of RPM analysis results[3]

The bi-directional shift factors allow us to plot master curves at any given reference temperature from data generate at multiple temperatures. The higher the test temperature, the more we accelerate the test.

3 Figure originally presented in final report for DOT PHMSA contract DTPH56-14-H-00001 project 554 https://primis.phmsa.dot.gov/matrix/PrjHome.rdm?prj=554&s=6D8E1B6104BE4DCEB642A4ACAF14CA84&c=1

Section 1 Bibliography

1. Kurdziel, J.M. and E.F. Palermo. *Stress Crack Resistance of Structural Members in Corrugated High-Density Polyethylene Pipe*. in *Transportation Research Board 86th Annual Meeting*. 2007.
2. ISO, *9080:2012 Plastics piping and ducting systems -- Determination of the long-term hydrostatic strength of thermoplastics materials in pipe form by extrapolation*. 2012, ISO.
3. Bragaw, C. *Service rating of polyethylene piping systems by the rate process method*. in *Proceedings Eighth Plastic Fuel Gas Pipe Symposium*. 1983.
4. Bragaw, C.G. *The Forecast of Polyethylene Pipe and Fitting Burst Life Using Rate Process Theory*. in *5th International Conference on Plastic Pipes, York UK*. 1982.
5. Coleman, B.D., *Application of the theory of absolute reaction rates to the creep failure of polymeric filaments.* Journal of Polymer Science, 1956. **20**(96): p. 447-455.
6. Montgomery, D.C. and G.C. Runger, *Applied statistics and probability for engineers*. 4th ed. 2007, Hoboken, NJ: Wiley. xvi, 768 p.
7. Lever, E., *Remaining Life Expectancy and Risk Profile of Vintage Aldyl A Piping*, in *Plastic Pipes XVII*. 2014: Chicago.
8. Lever, E., *Bi-Directional Shift Factors Revisited*, in *Plastic Pipes XVII*. 2014: Chicago, IL.
9. Mavridis, H. and R. Shroff, *Temperature dependence of polyolefin melt rheology.* Polymer Engineering & Science, 1992. **32**(23): p. 1778-1791.
10. Popelar, C. *A Comparison of the Rate Process Method and the Bidirectional Shifting Method*. in *Proceedings of the Thirteenth International Plastic Fuel Gas Pipe Symposium*. 1993.
11. Popelar, C., C. Popelar, and V. Kenner, *Viscoelastic material characterization and modeling for polyethylene.* Polymer Engineering & Science, 1990. **30**(10): p. 577-586.
12. Strobl, G.R., *The Physics of Polymers: Concepts for Understanding Their Structures and Behavior*. 2013: Springer Berlin Heidelberg.
13. Bower, D.I., *An Introduction to Polymer Physics*. 2002: Cambridge University Press.
14. Popelar, C.F., *Characterization of mechanical properties for polyethylene gas pipe materials*. 1989, The Ohio State University.

2. Stress Intensification Factors (SIF) for Aldyl-A Piping Systems

Brief Literature Review for Current Context

In **1** we have already seen that at its core, the Rate Process Method is a damage propagation model. Coleman [5] defines a variable, "γ", which he defines as a convenient measure of how much the applied stress has distorted the structure of the fiber at a time t. He is talking about the creep process that is governed by the α-relaxation of the polymer being evaluated. Bragaw [4], in addressing the burst life of polyethylene pipe and fittings, moves on to discuss the rupture process by defining an Arrhenius relationship between the fracture rate, the applied stress and the activation energy of the α-relaxation process in polyethylene. Neither Coleman, nor Bragaw, directly mention the α-relaxation process. However, from the literature [9, 11-13], and the coherence between the shift factors developed from the DTMA measurement of the α-relaxation and the RPM test results that are shifted to a reference temperature, it is clear that the α-relaxation process provides a good explanation of pipe rupture. Bragaw mentions that the Arrhenius equation only identifies a single activation energy, while there must clearly be many activated processes in a typical pipe rupturing process. Strobl [12], Strobl insert, covers this point adequately in noting that the activation energy for the α-relaxation process rolls up a large number of elementary processes. We now need to develop a better understanding of stress as a factor that drives variance in the calculated lifetime expectancy.

Failure Modes of Polyethylene

It is customary to identify three failure modes for polyethylene pipe as illustrated in **Figure 2-1.** In this project, we are interested in Region A and Region B as defined in the figure. The two failure regions are governed by the same α-relaxation process, only the degree of constraint at the damage tip is different in the two regions. The degree of constraint is discussed in detail in **8** below. For now, we can simply note that the magnitude of the parameters C2 and C4 in the RPM models shown in **Table 2-1** define the slope of each curve and reflect the degree of constraint at the damage tip. A cursory inspection of the equations will show that stress has a larger impact in Region A than in Region B. The shallowness of the slope in Region A reflects this. In Region A, the entire pipe wall is subject to the driving stress, whereas in Region B, a small volume at the damage tip is subject to the driving stress (the stress is highly constrained), and only this volume undergoes critical creep phenomena that are governed by the α-relaxation process (damage propagation through the wall – SCG).

Table 2-1. Aldyl A RPM models for Region A and Region B

Parameters	C1	C2	C4
Region B, Quasi-Brittle Failure	-17.6172	9485.337	-898.536
Region A, Ductile Failure	-42.5629	37387.00	-7515.74

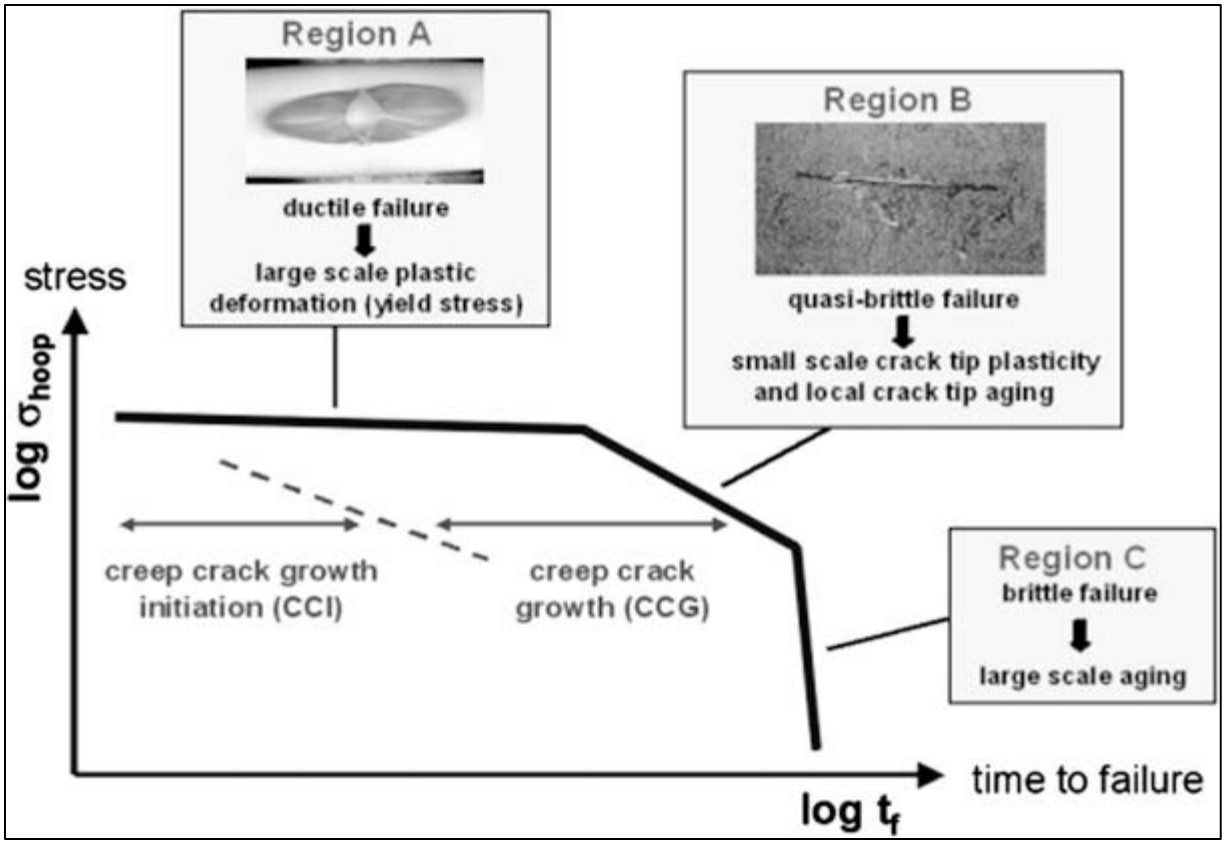


Figure 2-1. Schematic illustration of the failure behavior of pressurized PE pipes [15]

Variance of Stress in Pipes

Bragaw [4] begins his analysis from the fracture rate relationship defined in **Equation 2-1.**

$$\frac{dc}{dt} = \frac{KT}{\eta} e^{-[\frac{\varepsilon - \frac{\beta\sigma}{2}}{KT}]} \qquad \textbf{Equation 2-1}$$

dc/dt	Fracture rate
T	Absolute temperature in Kelvin
ε	Activation energy
σ	<u>True stress at activated sites</u>

Other symbols are constants

GTI adopts the ISO 9080 [2] formulation of the rate process model as defined in **Equation 2-2**.

$$\log(t) = C_1 + \frac{C_2}{T} + C_3\log(\sigma) + \frac{C_4\log(\sigma)}{T} \qquad \textbf{Equation 2-2}$$

C_1, C_2,C_3,C_4	Constants determined from data by linear regression (dependent on stress units used)
t	Time to failure in hours
T	Absolute temperature in Kelvin
σ	Stress usually in MPa, but can be in any stress unit. In this report psi is used to remain consistent with units commonly used in the industry in the USA.
log	Log base 10

The stress used in standard RPM analysis is the pipe hoop stress as derived from Barlow's formula that relates the internal pressure that a pipe can withstand to its dimensions and the strength of its material[4]

$$P = \frac{2St}{D}$$ **Equation 2-3**

P	Pressure
S	Allowable stress
T	Wall thickness
D	Outside diameter

For polyethylene pipe the formula is modified to take advantage of the practice of specifying the Dimension Ratio (DR) of the pipe, defined as the ratio between the pipe outside diameter and the wall thickness. If the DR is constrained to be a number from the Renard R10 series[5] it becomes a Standard Dimension Ratio (SDR). Pipe manufactured to a DR results in pipe with the same pressure capacity across all diameters under the assumption that the pipes are not thick-walled cylinders. This assumption holds for pipes of DR larger than 20 (radius/thickness<10), Young [16] § 13.2. We can immediately see that for the majority of

---

4 Barlow's formula. (2016, December 1). In Wikipedia, The Free Encyclopedia. Retrieved 11:28, September 20, 2017, from https://en.wikipedia.org/w/index.php?title=Barlow%27s_formula&oldid=752557552

5 Renard series are a system of preferred numbers dividing an interval from 1 to 10 into 5, 10, 20, or 40 steps. [1] This set of preferred numbers was proposed in the 1870s by French army engineer Colonel Charles Renard.[2] His system was adopted in 1952 as international standard ISO 3. Renard's system of preferred numbers divides the interval from 1 to 10 into 5, 10, 20, or 40 steps. The factor between two consecutive numbers in a Renard series is approximately constant (before rounding), namely the 5th, 10th, 20th, or 40th root of 10 (approximately 1.58, 1.26, 1.12, and 1.06, respectively), which leads to a geometric sequence. This way, the maximum relative error is minimized if an arbitrary number is replaced by the nearest Renard number multiplied by the appropriate power of 10.
Renard series. (2017, June 18). In Wikipedia, The Free Encyclopedia. Retrieved 11:34, September 20, 2017, from https://en.wikipedia.org/w/index.php?title=Renard_series&oldid=786314581

polyethylene piping applications the thin-walled cylinder assumption is not strictly true. The variance is not critical for determining the ductile-rupture boundary of the pipe as defined in PPI TR-3[6], however we should expect to see a significant component of variance in test results associated with the thin-walled assumption. The presence of stress-risers, or stress concentrations[7] will introduce additional variance.

---

6 http://plasticpipe.org/pdf/tr-3-2017a.pdf
TR-3/2017a HDB/HDS/PDB/ SDB/MRS/CRS Policies: Policies and Procedures for Developing Hydrostatic Design Basis (HDB), Hydrostatic Design Stresses (HDS), Pressure Design Basis (PDB), Strength Design Basis (SDB), Minimum Required Strength (MRS) Ratings, and Categorized Required Strength (CRS) for Thermoplastic Piping Materials or Pipe, **Section F.4.1.2**

7 A stress concentration (often called stress raisers or stress risers) is a location in an object where stress is concentrated. An object is stronger when force is evenly distributed over its area, so a reduction in area, e.g., caused by a crack, results in a localized increase in stress. A material can fail, via a propagating crack, when a concentrated stress exceeds the material's theoretical cohesive strength. The real fracture strength of a material is always lower than the theoretical value because most materials contain small cracks or contaminants (especially foreign particles) that concentrate stress. Fatigue cracks always start at stress raisers, so removing such defects increases the fatigue strength.
Stress concentration. (2017, July 1). In Wikipedia, The Free Encyclopedia. Retrieved 12:05, September 20, 2017, from
https://en.wikipedia.org/w/index.php?title=Stress_concentration&oldid=788498711

Stress Intensification Factors (SIF)

We will now develop basic SIF for the ductile failure mode of Aldyl A pipe following a process similar to that used in ASME B31.3 as described by Becht [17] §8.4. We do not use fatigue testing to determine the number of cycles to failure and then calculate a ratio to a reference configuration as is done in the ASME context, but do assume that the true stress in the pipe specimen is related to a stress intensification factor as is used in assessing damage tolerance:

> *"The **stress intensity factor**, , is used in fracture mechanics to predict the stress state ("stress intensity") near the tip of a crack caused by a remote load or residual stresses.[1] It is a theoretical construct usually applied to a homogeneous, linear elastic material and is useful for providing a failure criterion for brittle materials, and is a critical technique in the discipline of damage tolerance. The concept can also be applied to materials that exhibit small-scale yielding at a crack tip.*
> *The magnitude of depends on sample geometry, the size and location of the crack, and the magnitude and the modal distribution of loads on the material"*
>
> Stress intensity factor. (2017, September 20). In Wikipedia, The Free Encyclopedia. Retrieved 12:22, September 20, 2017, from https://en.wikipedia.org/w/index.php?title=Stress_intensity_factor&oldid=801550951

We can empirically estimate the SIF associated with each data point in a robust data set i.e. developed using consistent testing methodology and a statistically significant number of results that can be fitted to a model.

Assumptions:

1. All the variance in the data set is due to variance in the true stress at the point of failure (data sets with a known material variance can be used to develop an equivalent SIF for the known material condition e.g. LDIW)

Method:

1. Fit a regression model to the data set.
2. Find the point with the best relative performance by shifting the mean regression line to pass through each point, observe the new intercept and select the point with the largest intercept as being that with the best relative performance.
3. Define the regression model that passes through the best performing point as the reference model.
4. Calculate the reference model stress for each data point by inserting the actual failure time for the data point into the derived equation for stress given failure time.
5. Divide the reference model stress by the actual failure stress and define this ratio as the individual SIF for the data point

Figure 2-2 shows the DuPont data sets used to calculate empirical SIF using the above methodology, or variants of the methodology. **Figure 2-3** and **Figure 2-4** show the results for ductile failures.

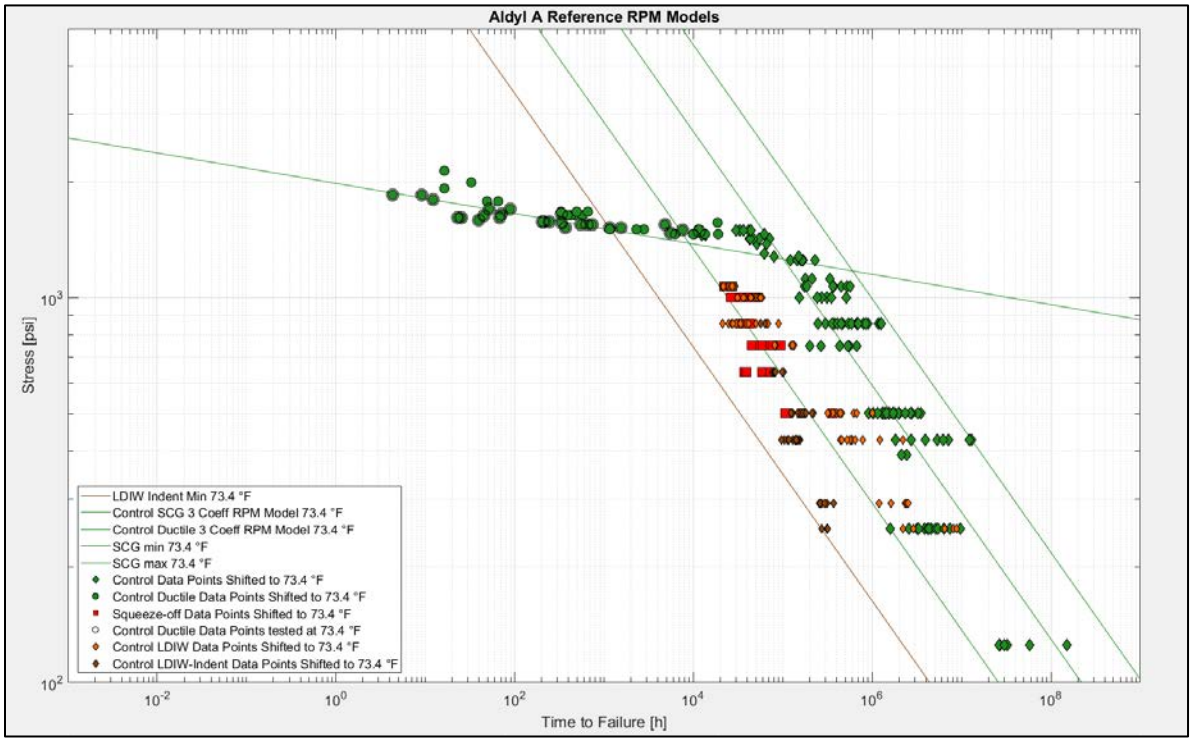


Figure 2-2. Aldyl A Data Sets for Deriving Empirical SIF

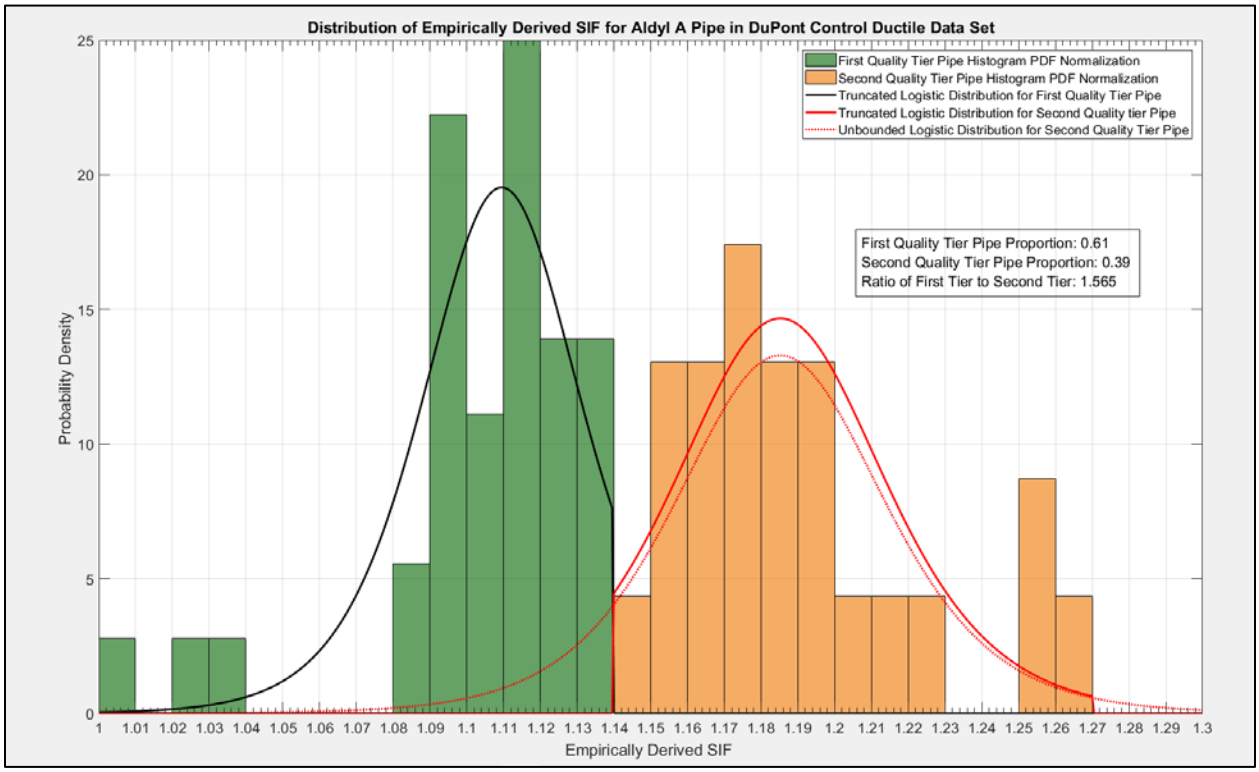


Figure 2-3. PDF for Empirically Derived SIF for Ductile Failures

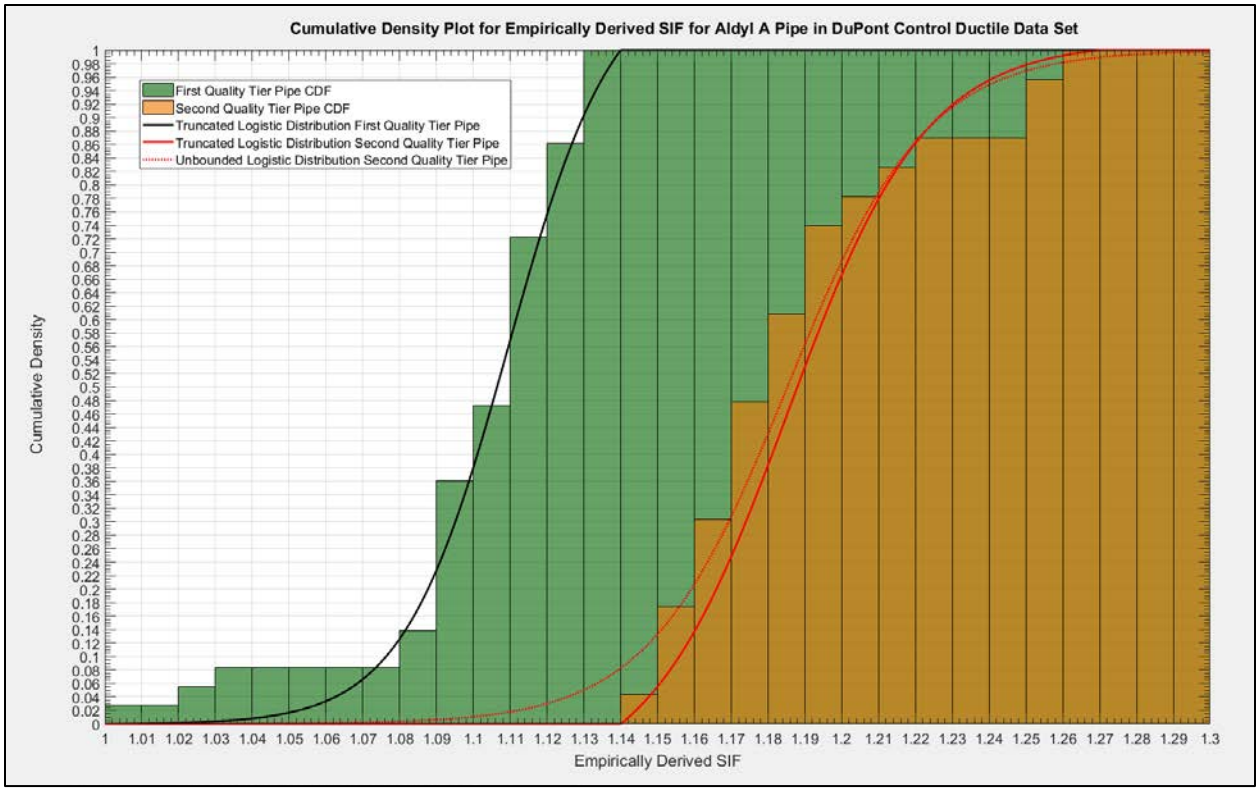


Figure 2-4. CDF for Empirically Derived SIF for Ductile Failures

Two distinct distributions of SIF are evident after the application of algorithm described above. It is reasonable to suggest that these two distributions reflect two quality levels of pipe as defined by the wall thickness irregularities, faceted inside diameter and die-lines always present to varying degrees in extruded pipe. To test this idea detailed FEM analyses were run to assess the impact of shallow grooves from die-lines on the SIF for the pipe as described below.

FEM analyses were run to determine the SIF due to sharp and blunt grooves introduced into the pipe internal diameter during extrusion. The assumption is that these SIF "seed" the ductile rupture process and that the ductile stress rupture curves reflect these SIF **Figure 2-5** and **Figure 2-6** show that at internal pressures typical of stress rupture tests the SIF for sharp grooves approaches 3, and for blunt grooves they approach 2.2. The SIF in the ductile data set was calculated as above, with the exception that the maximum SIF was set to 3 to match the maximum SIF from the FEM and the remaining SIF referenced to this value. The resulting distribution of ductile pipe SIF is shown in **Figure 2-7.** The minimum SIF calculates to 2.2, which is in remarkable agreement with the FEM analysis. This result is viewed as validation of the assumption that failure in the ductile regime is seeded by the axial scoring on the pipe ID.

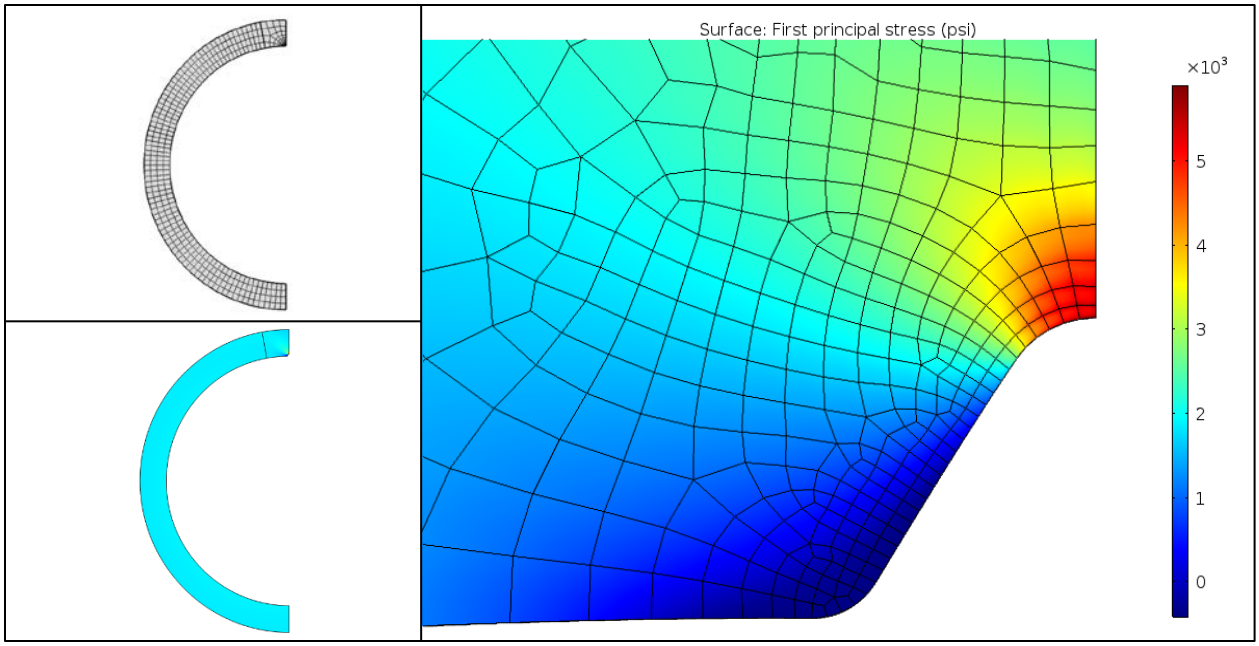


Figure 2-5. FEM analysis of stress associated with die line grooves on internal diameter of MDPE pipe

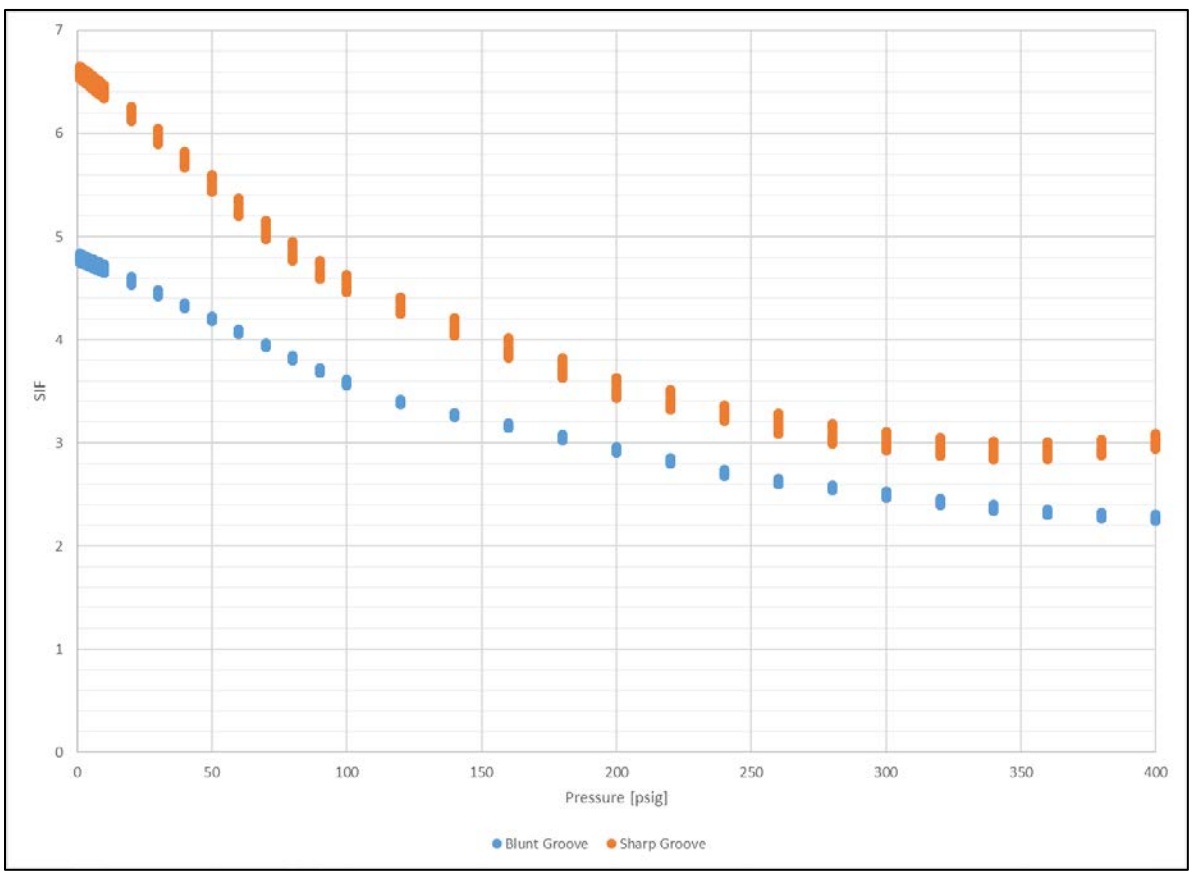


Figure 2-6. SIF measured in FEM as a function of internal pressure at 73.4°F

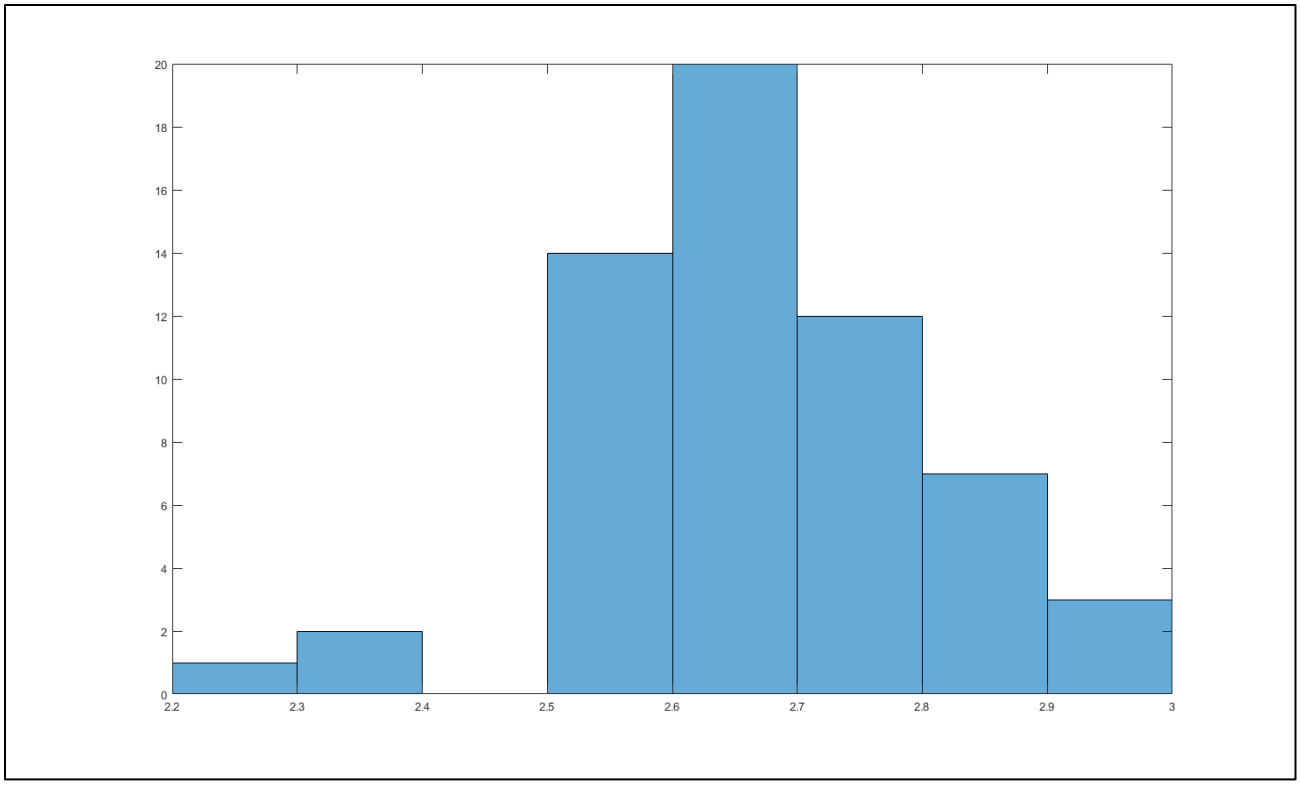


Figure 2-7. Distribution of ductile failure SIF referenced to FEM results

Important Caveat for Ductile Failure SIF

Actual testing of heavily scored pipe does not result in empirical SIF matching the FEM values, and the failure location does not match the groove locations. This is due to large scale plastic yielding of the pipe wall that reduces the SIF, and potentially the differential creep noted on page **329** that could lead to preferential yielding in the pipe wall resulting in eventual ductile rupture. The SIF distribution shown in **Figure 2-3** shows the effective SIF for different levels of pipe scoring because of the two processes described above playing out. Gas distribution pipe with deep grooves on the ID was extensively tested by GTI using the RPM **Figure 2-9**, and the results were referenced to baseline data from the resin supplier. The resulting SIF distributions are shown in **Figure 2-8**.

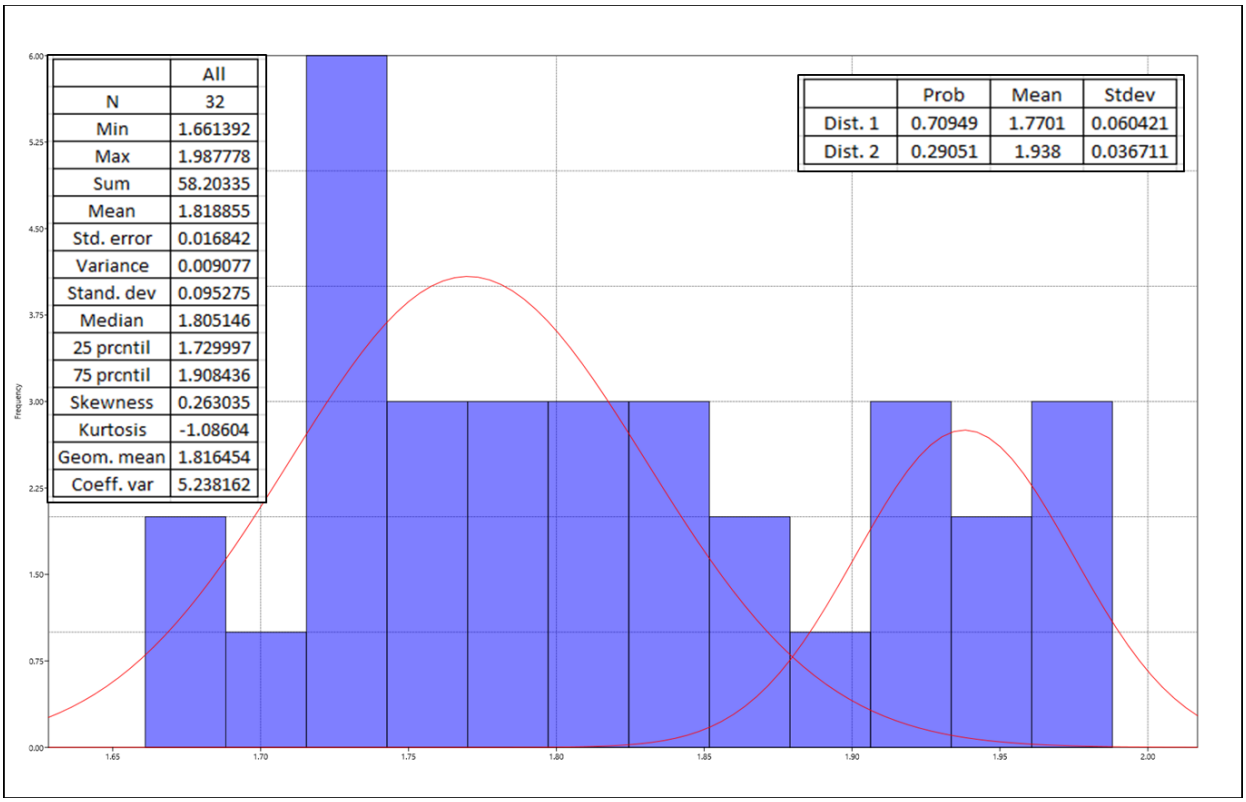


Figure 2-8. Empirical SIF Distributions for Heavily Grooved Gas Distribution Pipe

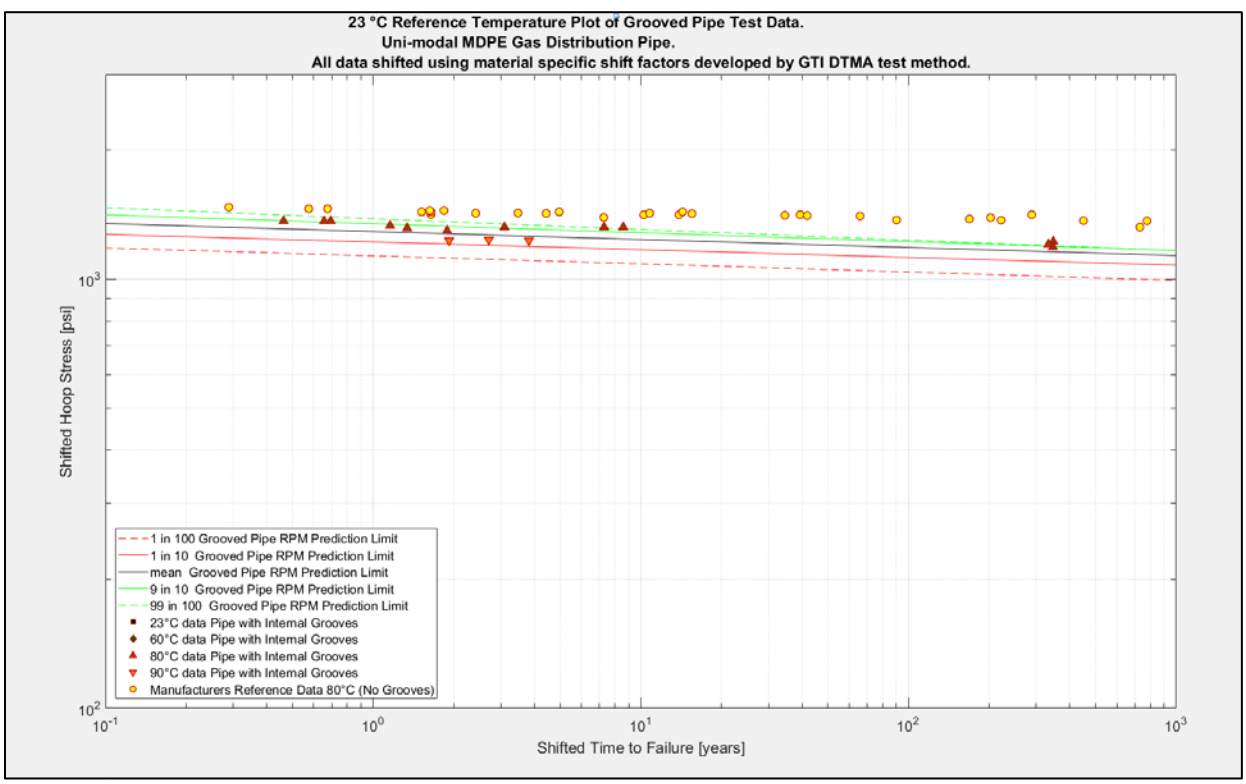


Figure 2-9. RPM Analysis for Empirical SIF Determination

The empirical method was extended to the SCG failures in the DuPont reference data sets and the resulting SIF distributions are shown in **Figure 2-10** together with the FEM adjusted ductile SIF distributions discussed above.

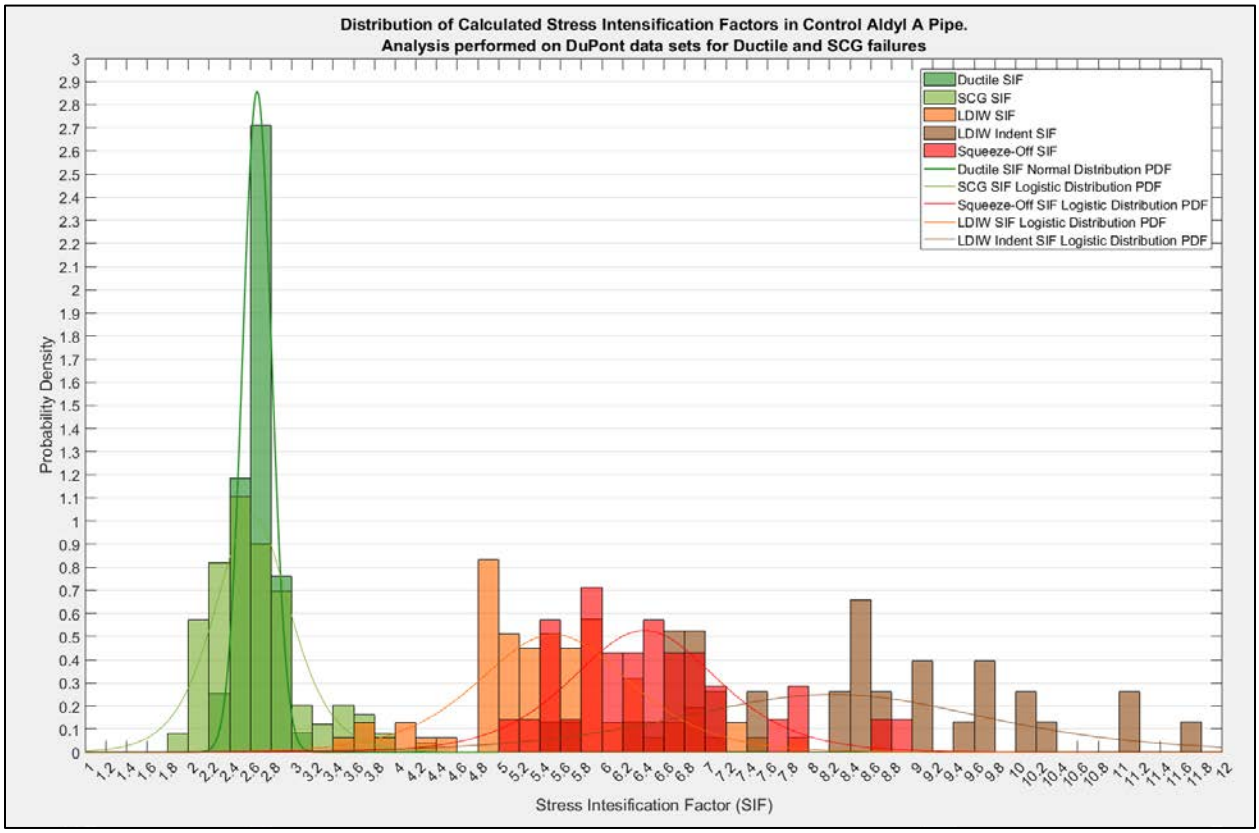


Figure 2-10. SIF Distributions for Various Aldyl A SCG Data Sets

The empirically derived SIF were used to predict expected lifetimes for the entire reference data set using the control DuPont SCG model with excellent results as shown in **Figure 2-11,** log-log plot, and **Figure 2-12,** linear-linear plot to emphasize variance. The variance in SCG failure times for a given stress over the performance range in **Figure 2-2** is a factor of 225**.** The SIF and reference model results yield a variance across the performance range of 2.5 as shown in **Figure 2-13** and **Figure 2-14.** That is, we get a two order of magnitude reduction in variance by addressing the true stress at the damage tip calculation via the empirically derived SIF, and using a single RPM reference model. This is a significant improvement over the simple hoop-stress approach and having to generate multiple RPM reference models.

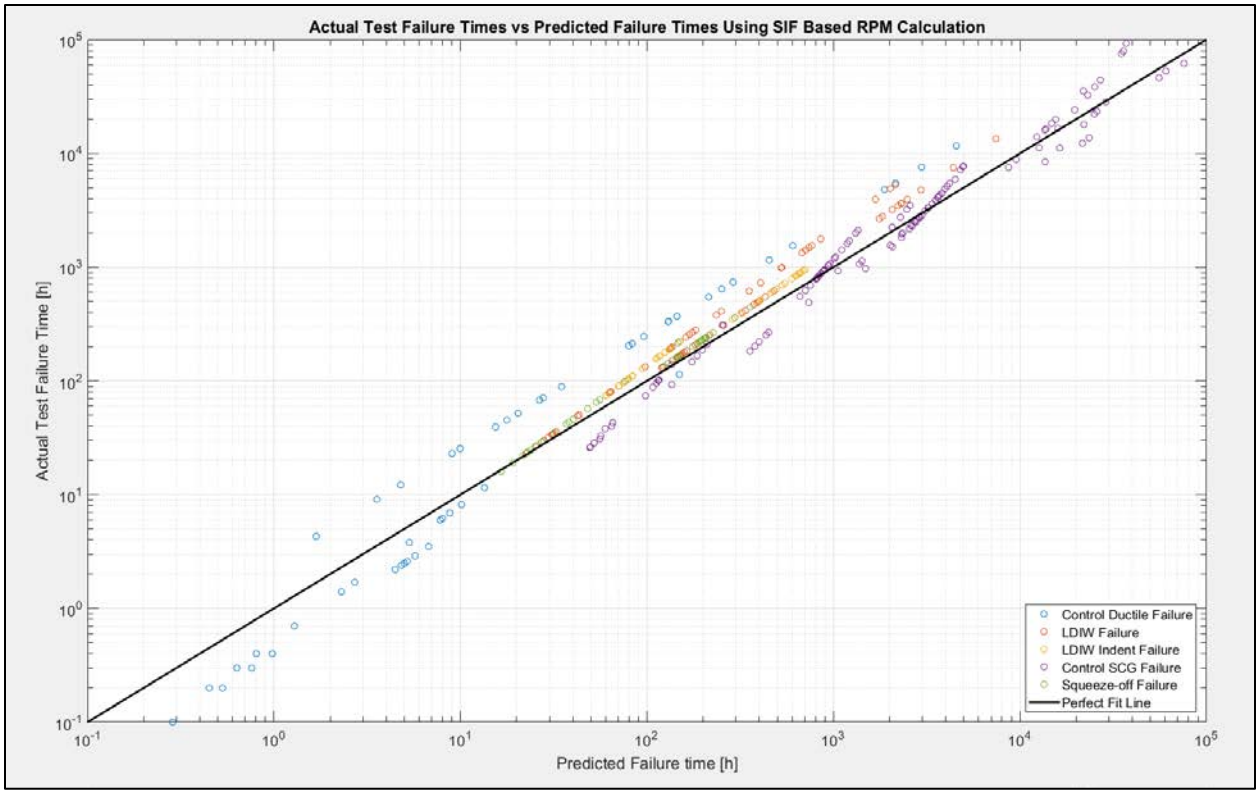


Figure 2-11. Actual vs Predicted Plot for SIF and Reference RPM Model Lifetime Prediction

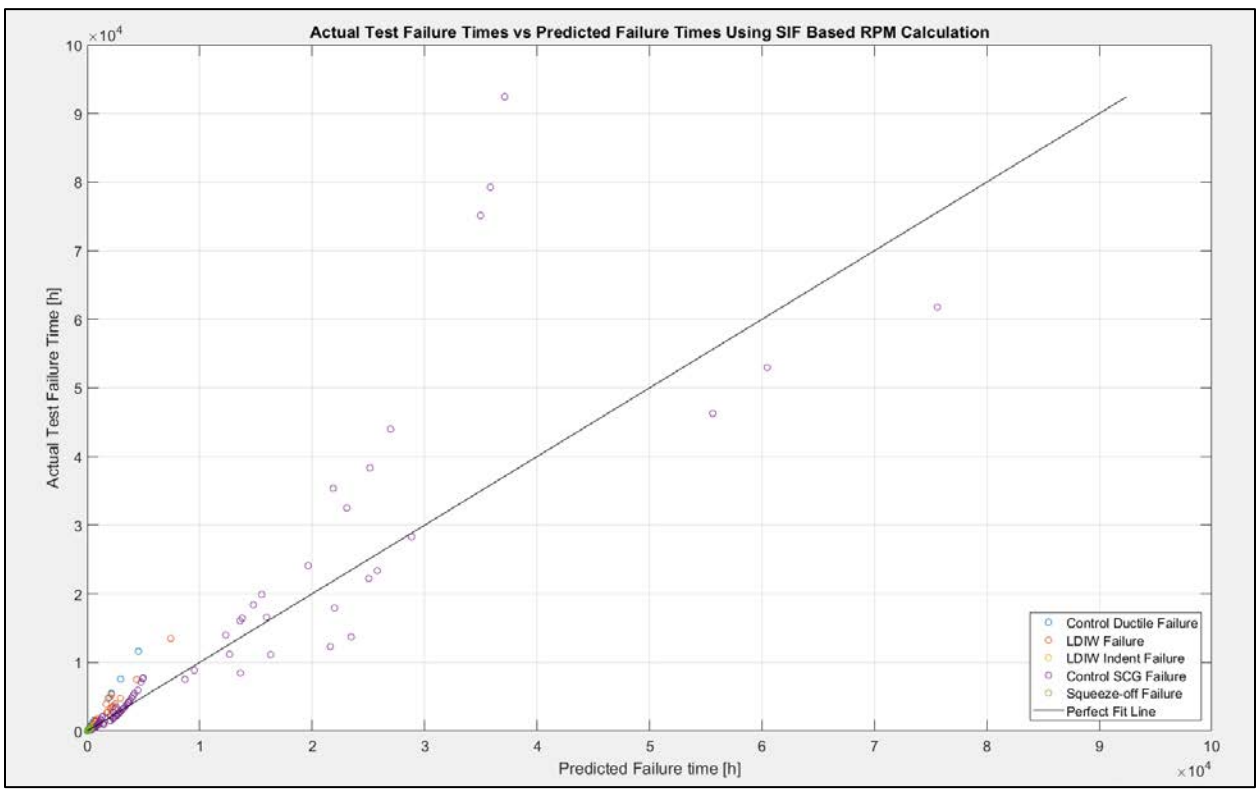


Figure 2-12. Same as Figure 2-11, but Linear Axes

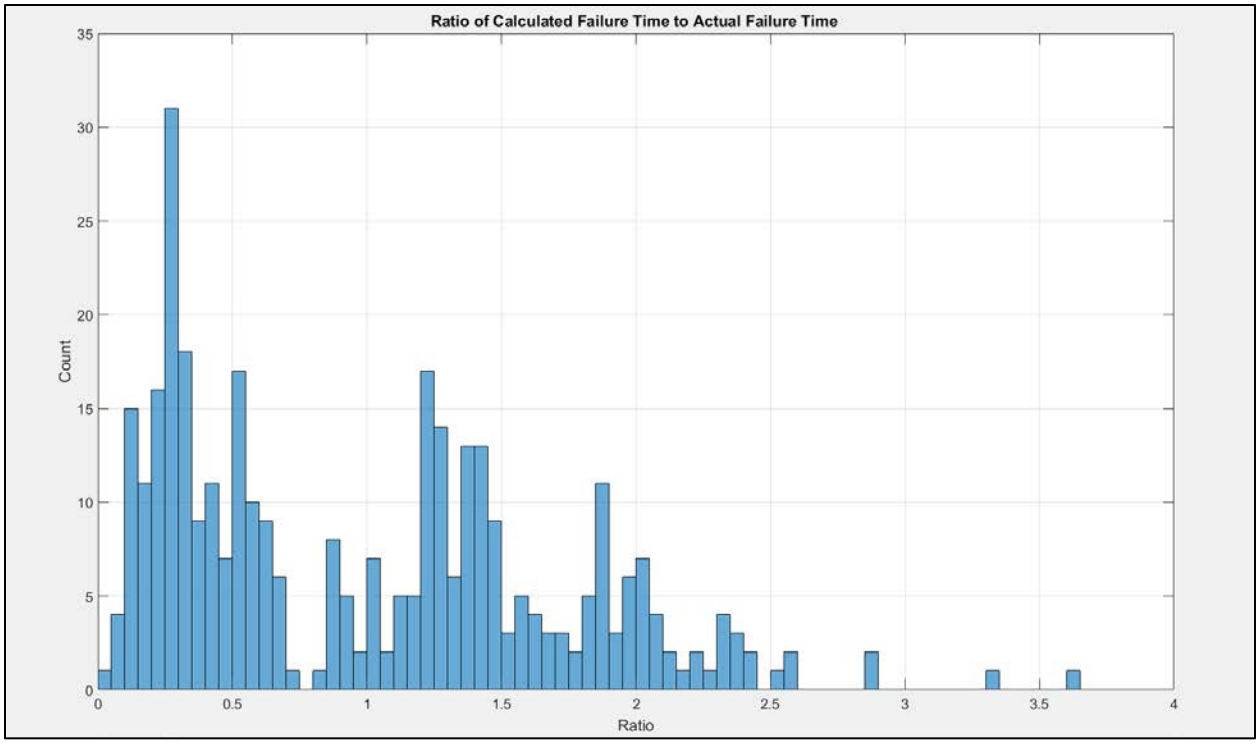


Figure 2-13. Actual to Predicted Failure Time Ratio

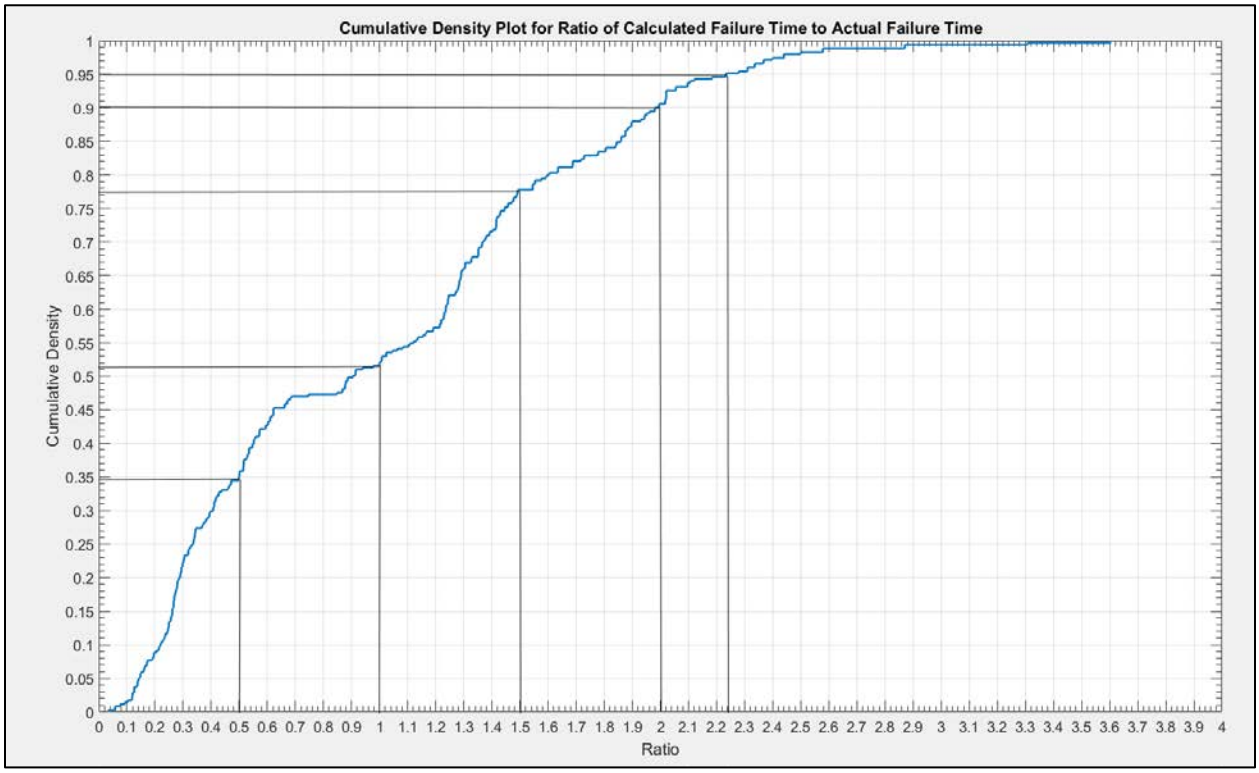


Figure 2-14. CDF for Actual to Predicted Failure Time Ratio

SIF for Polyethylene Piping Systems

In the preceding sections, we focused on extracting effective SIF from reference data sets under the assumption that the variance in performance implied by the results was entirely due to combinations of SIF in the test specimen. We will now look at SIF derived from FEM analyses of pipe and pipe assemblies. The FEM analyses all utilized advanced constitutive models for polyethylene based on extensive mechanical testing of multiple materials. The constitutive models incorporate temperature effects, strain rate effects and relaxation. The SIF listed below are all the SIF associated with initial loading of the assembly and represent the initial static stress state. **Table 2-2** lists some stand-alone SIF for different piping configurations. They are stand-alone because we have not yet addressed how to combine individual SIF into a composite effective SIF that can plugged into a RPM model.

Table 2-2. Stand-alone SIF for Polyethylene Piping Systems

Pipe or Fitting Configuration (All with 45 - 60 psig Internal Pressure, SDR11)	Stress Intensification Factor
Socket Coupling – Coupling Edge (FEA Analysis)	1.25
Soil Loading to 4% Deflection (FEA Analysis)	1.6
Saddle Tee (FEA Analysis)	2.7
Socket Coupling – Coupling Center (FEA Analysis)	1.8 – 2.9
Pipe with Bend Radius of 100 Pipe Diameters (FEA Analysis)	3.0
Bending (Empirical correlation from RPM testing)	3.4
Pipe with Bend Radius of 80 Pipe Diameters (FEA Analysis)	3.6
Pipe with Bend Radius of 50 Pipe Diameters (FEA Analysis)	4.7
Socket Coupling with Bend Radius of 100 Pipe Diameters (FEA Analysis)	4.8
Socket Coupling with Bend Radius of 80 Pipe Diameters (FEA Analysis)	5.8
Impingement (Empirical correlation from RPM testing)	5 - 7
Socket Coupling with Bend Radius of 50 Pipe Diameters (FEA Analysis)	7.5
Squeeze-Off at time of squeeze (FEA Analysis)	8.5 – 10.5

The SIF listed above correlate reasonably well with empirically derived SIF and give a plausible explanation for observed field failures and the proportions of individual causes reported by various sources. There is not much useful failure data in the public domain. Maupin and Mamoun [18] reviewed several hundred data points that included 55 field

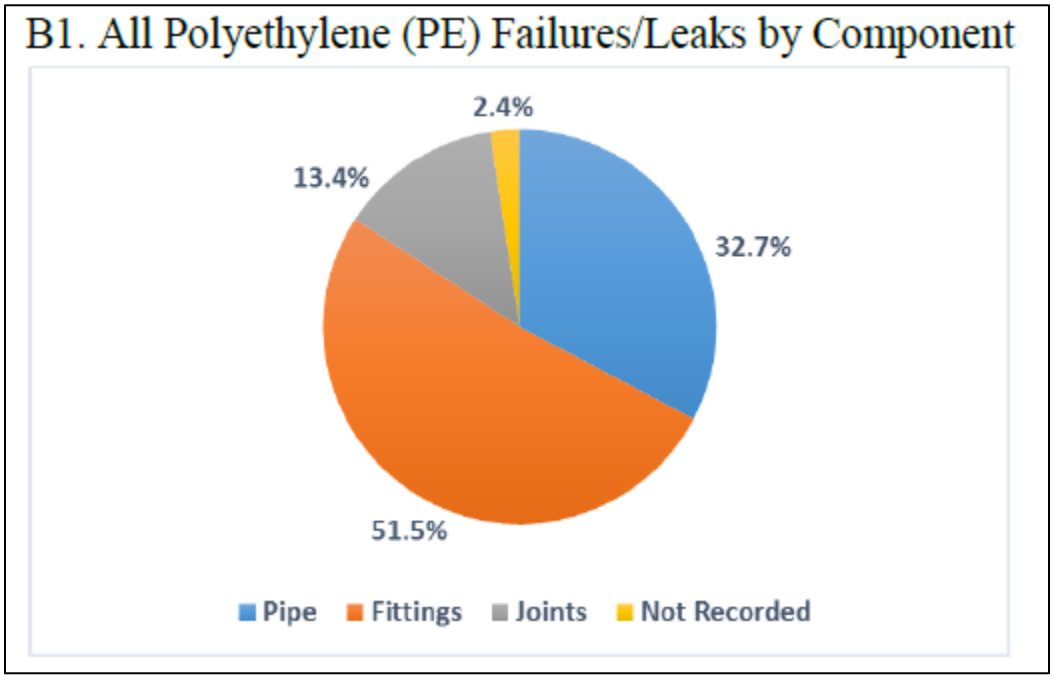
failures they analyzed, 104 failures in databases they had access to and 162 failure reports from other sources. Their summary of failure categories is presented in **Table 2-3.**

Table 2-3 Maupin and Mamoun summary of failure categories

Failure Type	Number
Material Failures	
Pipe	
Rock Impingement	9
Squeeze-off	8
Insert Renewal	1
Bending/Settlement	3
Internal Pressure	1
Joints	
End Caps	8
Tapping Tee Caps	9
Tees and Ells	21
Sockets	74
Saddles	118
Fusion Failures in Joints	
Butt Fusion	29
Socket Fusion	7
Saddle Fusion	5
Quality Control Problems	6
Third Party	14
Other	8
Total	**321**

Another source of information is the Plastic Pipe Database Collection Initiative (PPDC): "*A group of representatives of federal and state regulatory agencies and the natural gas and plastic pipe industries have come together and formed The Plastic Pipe Data Collection Initiative. Their goal has been to create a national database of information related to the in-service performance of plastic piping materials. Members include the American Gas Association, the American Public Gas Association, the Plastics Pipe Institute (PPI), the National Association of Regulatory Commissioners, the National Association of Pipeline Safety Representatives (NAPSR), the U.S. Department of Transportation and its Office of Pipeline Safety.*"
Source: https://www.aga.org/plastic-pipe-database-collection-initiative , accessed 9/21/2017

The PPDC April 2017 Status Report summarizes the data presented in **Figure 2-15**



All PE Failures/Leaks by Cause

CAUSE	% of All PE Failures/Leaks	% of All PE Pipe Failures/Leaks	% of All PE Fitting Failures/Leaks	% of All PE Joint Failures/Leaks
Excessive Expansion/Contraction	1.7%	1.2%	1.4%	4.3%
Excessive External Earth Loading	5.7%	9.0%	3.6%	4.8%
Installation Error	28.3%	12.2%	31.5%	56.4%
Squeeze Off	2.0%	5.7%	0.1%	0.1%
Point Loading	6.9%	16.2%	2.1%	3.0%
Previous Impact	1.9%	5.2%	0.3%	0.4%
Unknown	13.3%	9.1%	16.0%	13.6%
Other	14.0%	16.0%	14.5%	2.9%
Cap	4.9%	0.0%	9.2%	0.0%
Not Recorded	3.0%	2.9%	2.8%	3.7%
Material Defect	16.0%	18.8%	16.3%	10.3%
Gopher/rodent/worm damage	0.4%	1.2%	0.0%	0.0%
Unknown - Not Excavated, Replaced	1.2%	1.4%	0.6%	0.4%
Unknown - Abandoned	0.1%	0.1%	0.1%	0.0%
Corrosion	0.8%	0.8%	1.4%	0.1%
	100.0%	100.0%	100.0%	100.0%

Figure 2-15. PPDC April 2017 Data for Polyethylene

The only useful information we can glean from the PPDC data is that fittings dominate the failures, point loading, squeeze-off and earth loading appear as prominent causes. This information supports the relative severity of SIF presented in **Table 2-2.**

A more detailed FEM analysis of fittings subjected to bending was undertaken and it was found that saddle fittings represent one of the most severe combination of SIF for fittings in a polyethylene piping system. Studies, shown in **Figure 2-16,** were performed on three sizes of pipe 1.25",2",4", with and without tees. For each pipe configuration, the stress intensity factor (SIF) was calculated by taking the maximum von Mises stress in each load case and dividing it by the nominal pipe hoop stress calculated from the internal pressure of the respective load case. After comprehensive analysis, it is found that the calculated SIF is size independent, as expected, and all the regression models can be generalized to a single power law function that correlates the SIF with input configuration parameters, such as internal pressure (P) and bending radius factor (BRF) Equation **2-4**.

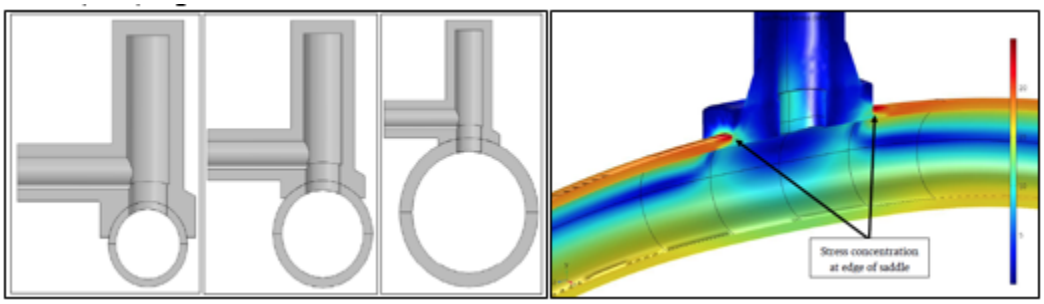


Figure 2-16. FEM models for saddle tee SIF evaluation

$$SIF = aP^b BRF^c$$ **Equation 2-4**

SIF – Stress Intensity Factor

P – Pressure[psi]

BRF – Bend Radius Factor expressed as multiples of the pipe diameter

a, b, c – Regression Coefficients

The FEM models were extended to consider the direction of the pipe bending relative to the fitting placement for saddle tees, which are not symmetric in this context. Socket couplings are symmetric and the direction of bending is immaterial. The bending moment sign convention used in this analysis is defined in **Figure 2-18.** The tee will be on the concave up portion of the pipe for positive bending, and on the convex up portion of the pipe for negative bending. The correlation coefficients for the key configurations are presented in **Table 2-4.** A plot of the SIF for a pipe without saddle tee is illustrated in **Figure 2-17**

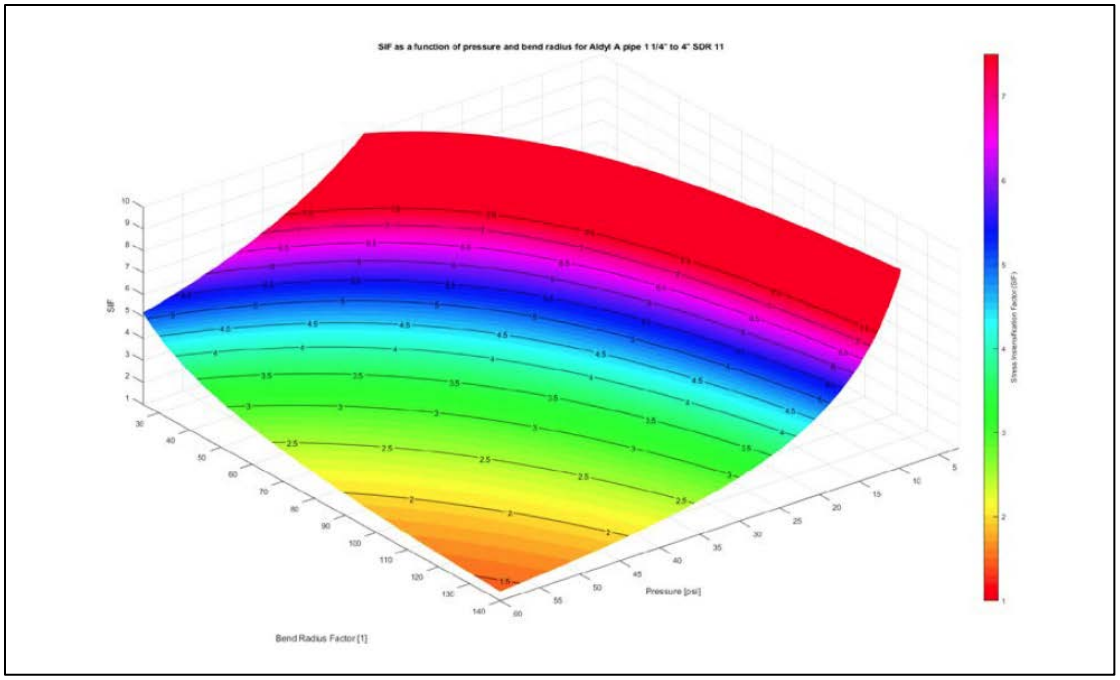


Figure 2-17. SIF for pipe as a function of BRF and P

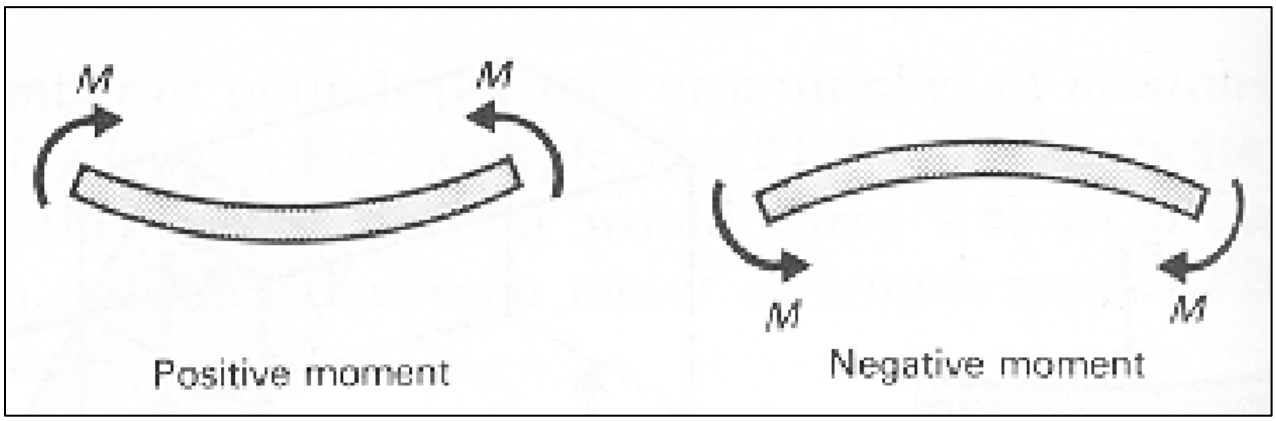


Figure 2-18. Bending moment sign convention-positive concave up, negative concave down

Table 2-4. Correlation coefficients for pipe under bending with and without fittings

Configuration	a	b	c
Pipe without Saddle Tee	4453.0	-1.0657	-0.744
Pipe with Saddle Tee Positive Bending	5310.7	-0.9654	-0.7251
Pipe with Saddle Tee Negative Bending	9364.1	-1.0757	-0.7789
Pipe with Saddle Tee Lateral	3476.7	-1.0059	-0.5593
Pipe with Coupling	4285.0	-0.676	-0.9252

The relative effect of the different bending configurations on saddle tee assemblies is presented in **Figure 2-19** and the SIF for pipe, tee and couplings in **Figure 2-20.**

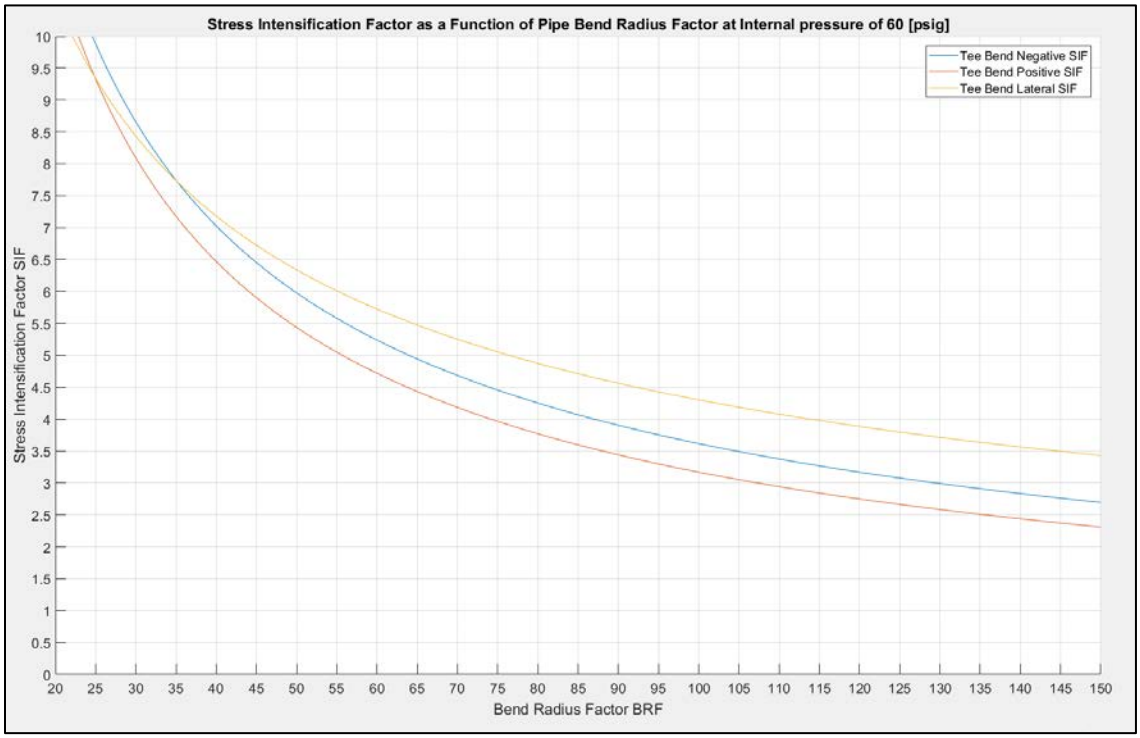


Figure 2-19. Relative effect of bending configurations on saddle tee assemblies

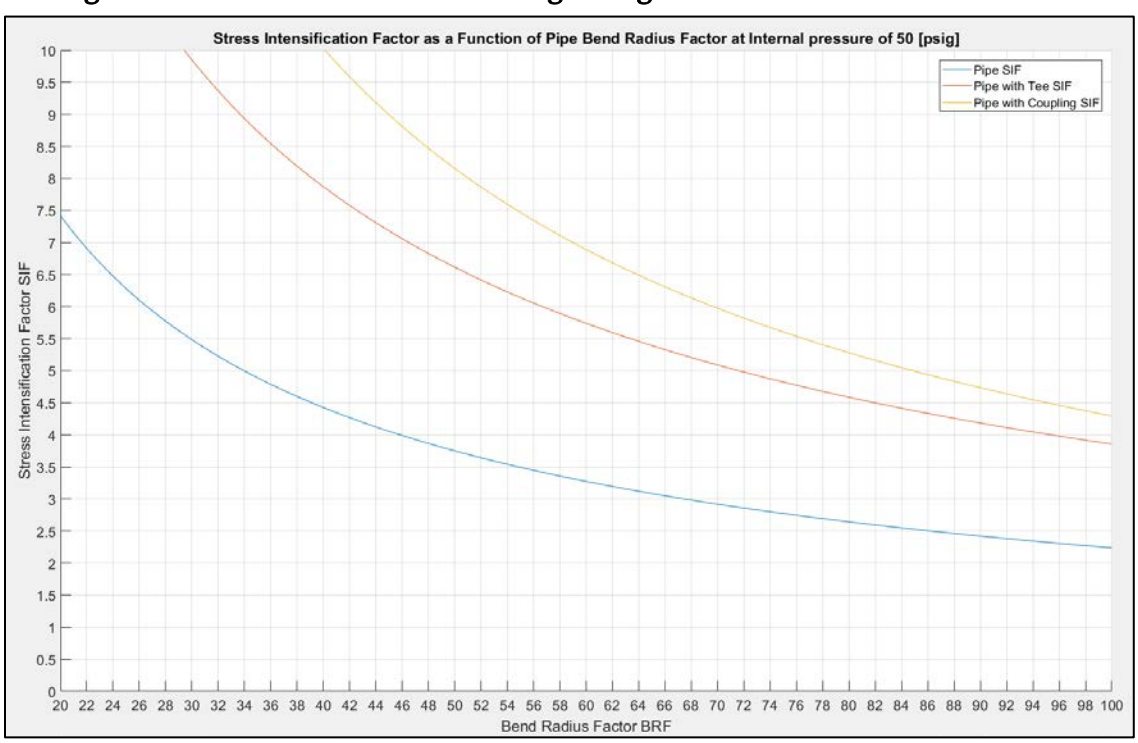


Figure 2-20. SIF for pipe, tee and coupling under bending

Combining SIF Distributions for Polyethylene Piping systems

In the sections above we have presented several SIF distributions extracted empirically from actual test data. Through inspection of many of these data sets, it was found that a Logistic Distribution provides the best fit for the data. The logistic distribution is a two-parameter distribution with a location parameter that is the mean of the variable being modeled and a shape factor that reflects the variance of the variable **Figure 2-21**.

Parameters	μ location (real) $s > 0$ scale (real)
Support	$x \in (-\infty, \infty)$
PDF	$\frac{e^{-\frac{x-\mu}{s}}}{s\left(1+e^{-\frac{x-\mu}{s}}\right)^2}$
CDF	$\frac{1}{1+e^{-\frac{x-\mu}{s}}}$
Mean	μ
Median	μ
Mode	μ
Variance	$\frac{s^2\pi^2}{3}$
Skewness	0
Ex. kurtosis	1.2
Entropy	$\ln(s)+2=\ln(\sigma)+1.40457\epsilon$ where σ is the standard deviation.
MGF	$e^{\mu t}\,\mathrm{B}(1-st, 1+st$ for $st \in (-1, 1)$ Beta function
CF	$e^{it\mu}\frac{\pi st}{\sinh(\pi st)}$

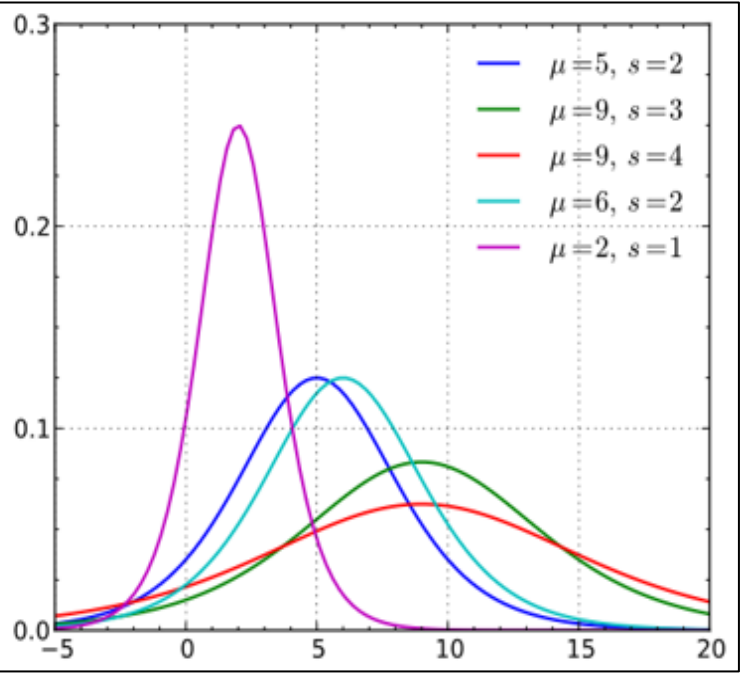


Figure 2-21. Wikipedia contributors. "Logistic distribution." *Wikipedia, The Free Encyclopedia*. Wikipedia, The Free Encyclopedia, 15 Jul. 2017. Web.

The logistic distribution resembles the normal distribution, but has heavier tails making it better suited to simulation techniques where there is uncertainty as to the actual values of the parameters.

Inspection of several empirically derived logistic distributions for SIF in polyethylene pipe systems leads to the "rule of thumb" that setting the shape factor "s" to 6 percent of the mean value is a reasonable first approximation that can be updated as more data is gathered. It was also found that a convenient approach to combining multiple SIF acting on a single piping component is to use the Euclidean, or L2 norm to sum them. We can represent "n" independent SIF as a vector:

$$\mathrm{SIF} = (\mathrm{SIF1}, \mathrm{SIF2}, \ldots . \mathrm{SIFn})$$

The magnitude of the composite SIF is given by **Equation 2-5**

$$||\mathrm{SIF}||_2 = \sqrt{\mathrm{SIF}_1^2 + \cdots + \mathrm{SIF}_n^2} \qquad \textbf{Equation 2-5}$$

We can use Equation 2-5 to reference all SIF to a baseline SIF that we define as Pipe Quality Tier 1 as discussed above and shown in **Figure 2-3** by requiring the L2 norm of every pair of SIF that includes the baseline SIF to be equal to the standalone value of the second SIF. The same method can be applied to the shape factors. **Table 2-5** presents a collection of SIF for both the ductile and SCG failure modes, each group being referenced to the Pipe Quality Tier 1 SIF for the relevant failure mode. SIF for Ductile and SCG failure modes need to be treated separately. For the most part we are only interested in the SCG failure mode.

Table 2-5. SIF Distributions for Polyethylene Piping systems

SIF for Ductile Failure: Logistic Distribution Parameters	Standalone SIF		Referenced to baseline pipe SIF	
	Mu	sigma	Mu	sigma
Pipe Quality Tier 1 – Baseline SIF	1.110	0.014	1.110	0.014
Use standalone SIF value for single SIF. Develop composite SIF by adding SIF referenced to baseline using Equation 2-5 for each independent SIF component acting on pipe component to the baseline SIF				
Pipe Quality Tier 2	1.185	0.018	0.417	0.011
Pipe Quality Tier 3 Severe Grooves	1.819	0.109	1.441	0.109
SIF for SCG Failure: Logistic Distribution Parameters	**Standalone SIF**		**Relative to baseline pipe SIF**	
	Mu	**sigma**	**Mu**	**sigma**
Pipe Quality Tier 1 SCG – Baseline SIF	2.000	0.120	2.000	0.120
Use standalone SIF value for single SIF. Develop composite SIF by adding SIF referenced to baseline using Equation 2-5 for each independent SIF component acting on pipe component to the baseline SIF				
Pipe Quality Tier 2 SCG	2.200	0.132	0.917	0.055
Pipe Quality Tier 3 Severe Grooves SCG	3.000	0.180	2.236	0.134
Pipe Quality Tier 4 LDIW SCG	5.780	0.347	5.423	0.326
LDIW Squeeze-off SCG	6.696	0.402	6.390	0.384
LDIW Impingement SCG	8.598	0.516	8.362	0.502

SIF Summary

- We have discussed the failure mechanism of polyethylene in some detail
- We have discussed the variance of stress in polyethylene pipes
- We have defined our usage of the term SIF
- We have developed SIF for various polyethylene piping configurations
- We have shown how to develop a probability distribution for SIF using a mean value and a quantification of our uncertainty for the mean
- We have presented a method for combining multiple SIF in a coherent manner

Section 2 Bibliography

2. ISO, *9080:2012 Plastics piping and ducting systems -- Determination of the long-term hydrostatic strength of thermoplastics materials in pipe form by extrapolation*. 2012, ISO.
4. Bragaw, C.G. *The Forecast of Polyethylene Pipe and Fitting Burst Life Using Rate Process Theory*. in *5th International Conference on Plastic Pipes, York UK*. 1982.
5. Coleman, B.D., *Application of the theory of absolute reaction rates to the creep failure of polymeric filaments.* Journal of Polymer Science, 1956. **20**(96): p. 447-455.
9. Mavridis, H. and R. Shroff, *Temperature dependence of polyolefin melt rheology.* Polymer Engineering & Science, 1992. **32**(23): p. 1778-1791.
11. Popelar, C., C. Popelar, and V. Kenner, *Viscoelastic material characterization and modeling for polyethylene.* Polymer Engineering & Science, 1990. **30**(10): p. 577-586.
12. Strobl, G.R., *The Physics of Polymers: Concepts for Understanding Their Structures and Behavior*. 2013: Springer Berlin Heidelberg.
13. Bower, D.I., *An Introduction to Polymer Physics*. 2002: Cambridge University Press.
15. Frank, A., G. Pinter, and R.W. Lang, *Prediction of the remaining lifetime of polyethylene pipes after up to 30 years in use.* Polymer Testing¶, 2009. **28¶**(7¶): p. 737-745.
16. Young, W.C. and R.G. Budynas, *Roark's Formulas for Stress and Strain*. 2001: McGraw-Hill Education.
17. Becht, C., *Process Piping: The Complete Guide to ASME B31.3*. 2009: ASME Press.
18. Maupin, J.M., Michael, *DTPH56-06-T-0004: Plastic Pipe Failure, Risk,and Threat Analysis*. 2009, GTI.

3. Adjusted RPM Models for Use with Aldyl-A SIF

So far, we have established a sound reference RPM model for Aldyl A that can be used to predicted the lifetime of a piping component, provided we have an accurate representation of the stress driving the mode of failure. We have developed SIF to represent the driving stress, provided a means for combining SIF and are left with the task of reconciling the RPM model with the SIF.

The complicating issue is that the RPM models were developed with test specimens having a wide variance of actual effective-SIF during their evaluation. In the section **Stress Intensification Factors (SIF)** above, we presented a method for estimating the distribution of these SIF in a dataset under the assumption that the variance in test results is due solely to the distribution of effective SIF. We showed that the standalone SIF we developed match SIF determined by various other methods quite closely.

Adjusting the RPM Model for Internal SIF

We will now adjust the RPM model for SCG in Aldyl A to account for the SIF that were present in the test specimens, normalize the standalone SIF accordingly, and demonstrate the validity of the method using several independent datasets.

Figure 3-1 shows the distribution of SIF in the DuPont control dataset for SCG failures. It is interesting to note that the mean SIF value is very close to that shown in **Figure 2-6** for blunt grooves, and that there is peak around SIF=3 that would correspond to sharp grooves. We would expect to realize the full effect of SIF in SCG type failures. **Figure 3-2** shows the true stress in the test specimens, obtained by multiplying the hoop stress by the SIF developed from assuming that the variance in results is solely due to differences in the true stress.

On substituting the true-stress values into a RPM analyses we find that 94 of the 122 data points closely follow the expected Arrhenius behavior as evidenced by a straight line on a log-log plot. Twenty-nine data points deviate markedly from this behavior. This is shown in **Figure 3-3.** In selecting data points to include in the model the 94 well behaved points were selected by default and a further 5 points were added to ensure that the mean line of the RPM model closely matched the slope of the 94 points on the log-log plot. The justification for this step is that the slope of the model lines reflects a combination of the activation energy for the α-relaxation process and the degree of constraint at the damage tip. The data points that deviate from this behavior may have been mixed mode failures where the degree of constraint at the damage tip was greatly reduced, or they could reflect data recording or other experimental errors. The resulting adjusted RPM model is presented in **Table 3-1**.

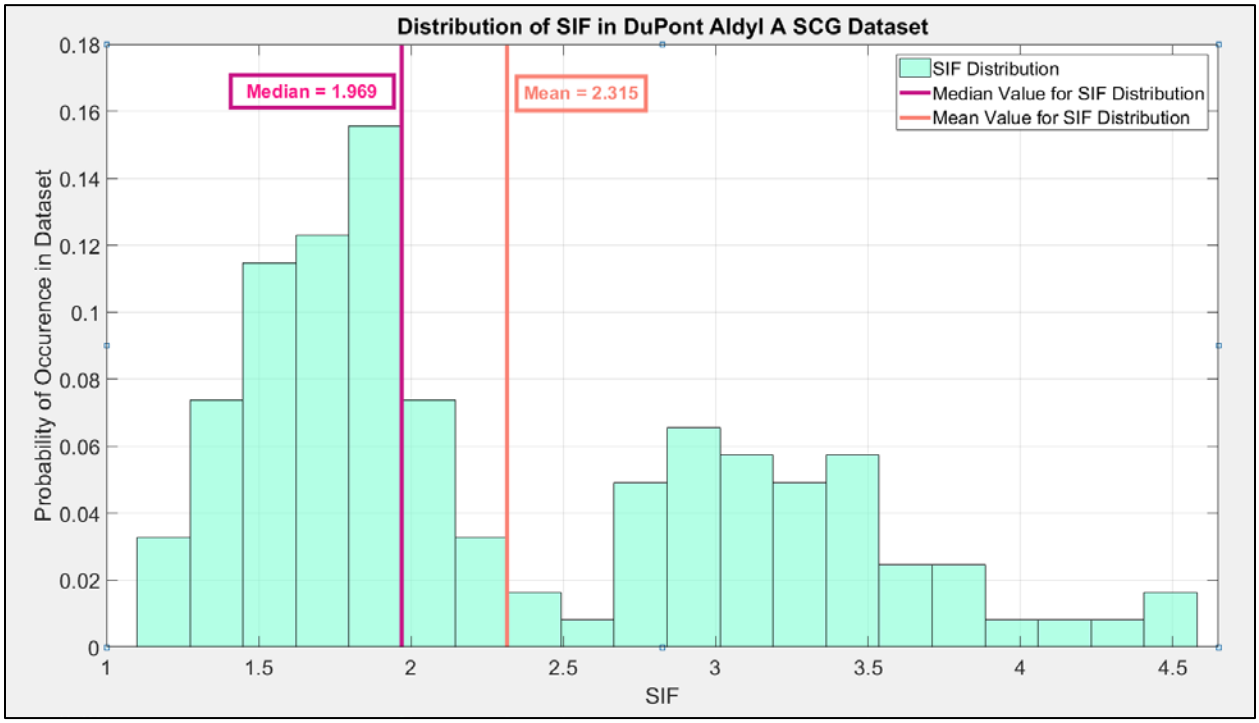


Figure 3-1. Distribution of SIF in DuPont SCG Reference Dataset

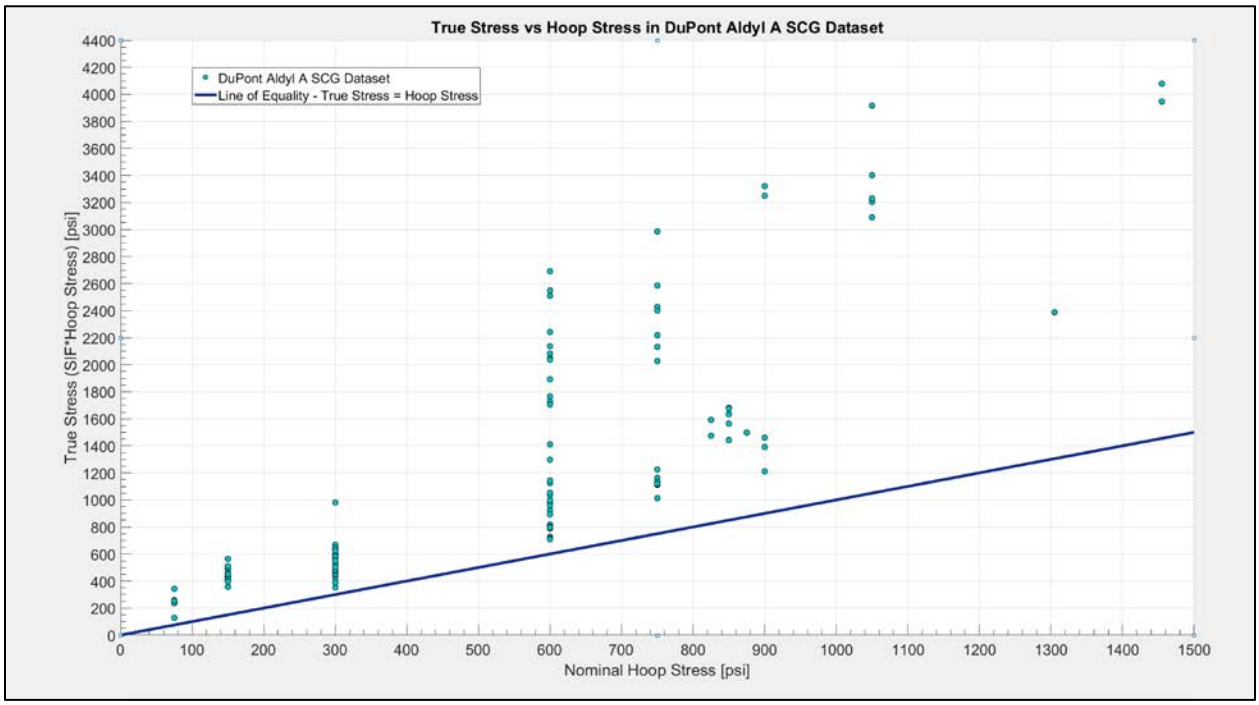


Figure 3-2. True Stress vs Hoop Stress in DuPont SCG Reference Set

Table 3-1. SIF Adjusted DuPont SCG Control Model Parameters for 3 Parameter ISO 9080 Model

Parameters	C1	C2	C4	R^2	R^2_{adj}	n	p	σ^2
Value	-18.051	10268.4	-1030.7	0.954	0.953	99	3	0.0429
Standard Error	0.52565	232.21	28.0683					

Covariance Matrix		
0.276312393	-116.2678638	7.443808708
-116.2678638	53921.28778	-4836.675907
7.443808708	-4836.675907	787.830244

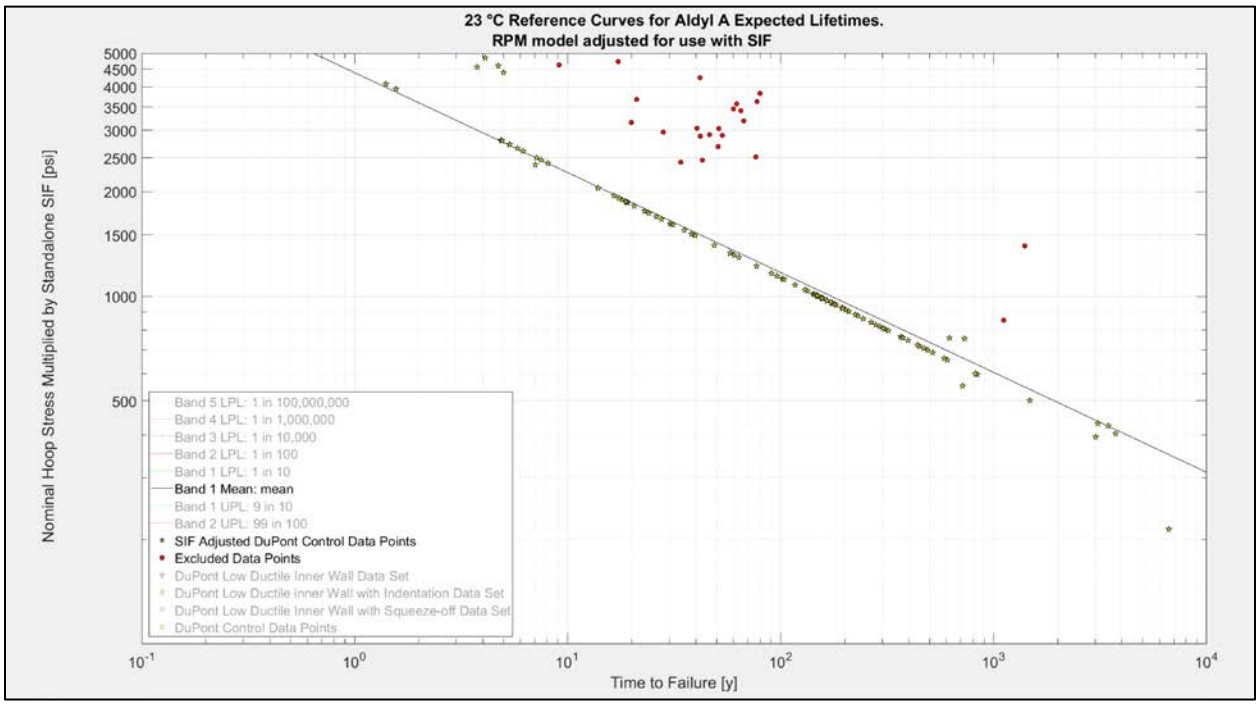


Figure 3-3. SIF Adjusted RPM model for DuPont Control SCG Dataset

The prediction limit bands for the adjusted RPM model are shown in **Figure 3-4**. There is no intrinsic significance to the prediction bands as they do not reflect true variance in a dataset, they simply give a convenient measure of distance from the mean for when we plot actual data sets for comparison.

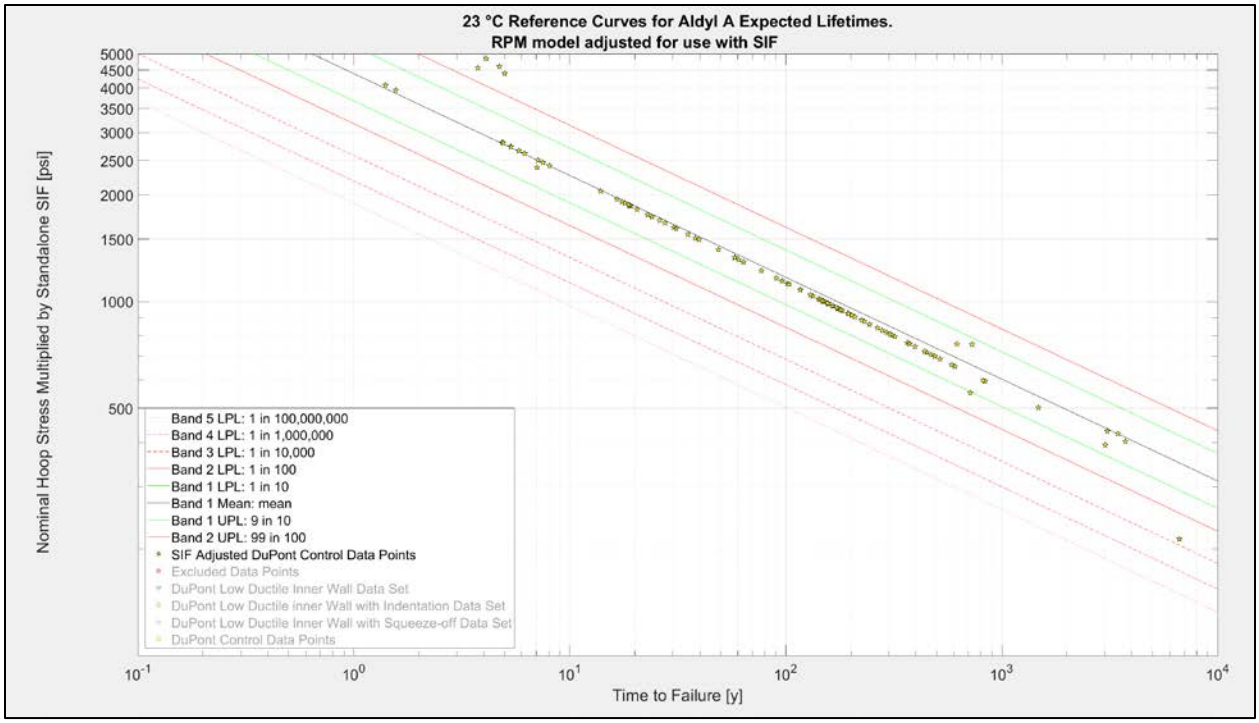


Figure 3-4. Adjusted RPM Model with Prediction Limit Bands

In addition to the control SCG dataset GTI has DuPont reference datasets for Aldyl A pipe with LDIW, LDIW with Squeeze-off, and LDIW with indentation. These four datasets are independent and reflect varying levels of intrinsic defect and material condition. The hoop stress for each data set was multiplied by the standalone SIF presented in **Table 3-2**. The SIF adjusted for use with adjusted RPM model are normalized to the median of the SIF for the control SCG dataset.

Table 3-2. Adjusted SIF for use with Adjusted RPM Model

Adjusted SIF for SCG Failure: Logistic Distribution Parameters for use with adjusted RPM model given in Table 3-1	Standalone SIF		Relative to baseline pipe SIF	
	Mu	sigma	Mu	sigma
Pipe Quality Tier 1 SCG – Baseline SIF	1.016	0.061	1.016	0.061
Use standalone SIF value for single SIF. Develop composite SIF by adding SIF referenced to baseline using Equation 2-5 for each independent SIF component acting on pipe component to the baseline SIF				
Pipe Quality Tier 2 SCG	1.117	0.067	0.461	0.028
Pipe Quality Tier 3 Severe Grooves SCG	1.524	0.091	1.136	0.068
Pipe Quality Tier 4 LDIW SCG	2.936	0.176	2.755	0.165
LDIW Squeeze-off SCG	3.401	0.204	3.246	0.195
LDIW Impingement SCG	4.367	0.262	4.247	0.255

The SIF adjusted data points for the four datasets are overlaid on the adjusted RPM model in **Figure 3-5**.

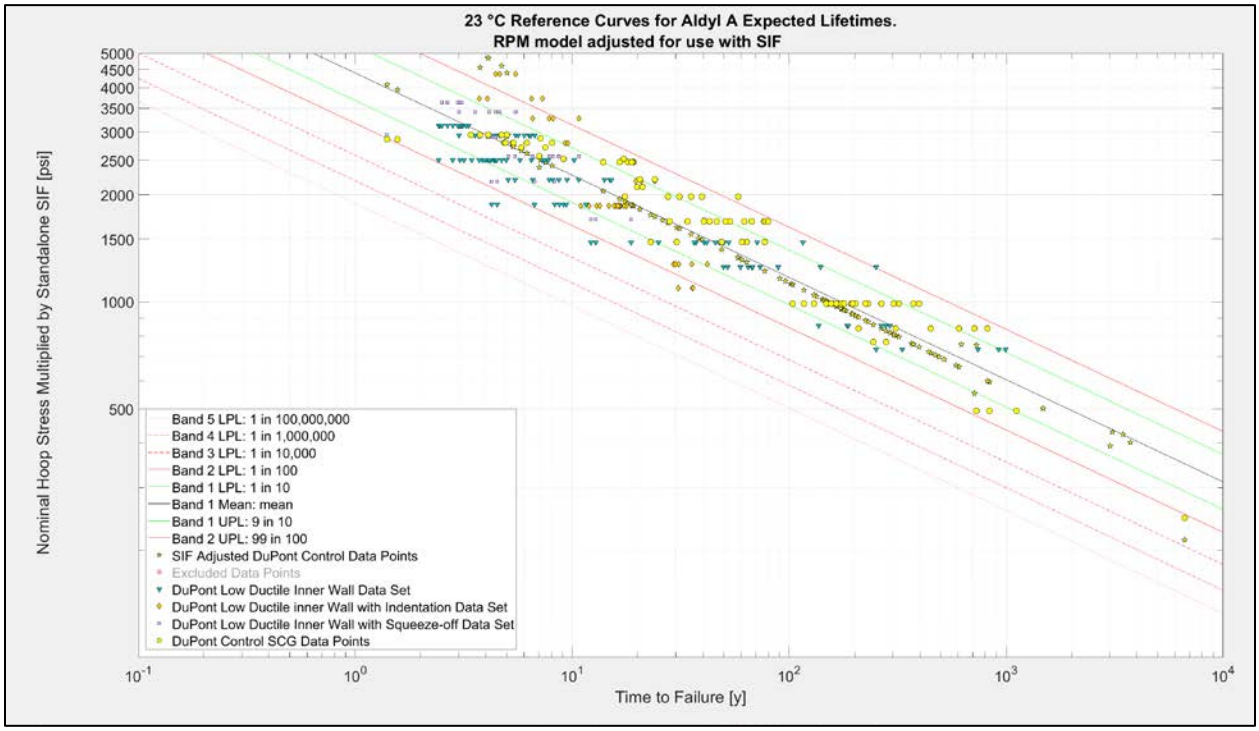


Figure 3-5. Validation of Adjusted RPM Model and Standalone SIF Approach for 4 Independent Datasets with Different Intrinsic Defects and Loading Conditions

All four of the datasets overlap in the same region and capture the same slope as the model indicating the adjusted model is a true material behavior model calibrated to the true stress at the damage tip. There are 7 points out of 351 above the Band 2 UPL, or 1.02% and 8 points below the Band 2 LPL, or 1.17%. We would be expecting to find 1% of data points above and a further 1% of data points below these two prediction bounds respectively. Given this excellent agreement we can say that the proposed methodology gives us 98% certainty in our predictions.

Summary

1. We have developed a RPM prediction model for Aldyl A that captures the SCG behavior if we know the true stress at the damage tip
2. We have developed SIF for many known material conditions and loading configurations that are calibrated to the RPM model
3. We have demonstrated the validity of the RPM model and SIF with four independent datasets containing 351 validation points.
4. We have demonstrated that the prediction limits of the RPM model match the reference data to the expected confidence level.

4. Bayesian Network Model for Aldyl-A Lifetime Expectancy

Introduction

In this section, the roots causes for pipe risk assessment are identified and an ontology describing the interactions between them is proposed, which primarily includes the current knowledge base of the subject matter experts. Upon that, a directed acyclic graph (DAC) approach, i.e. Bayesian network is employed to express the causal relationship between the root causes. Different from the classical index-based risk model, the probabilistic Bayesian network model can inherently incorporate the historic data, subject matter expert's opinions, as well as the belief about the collected data. Those characteristics allow for continuous refinement of the network structure and node probability table when additional knowledge is available. The risk prognosis and context condition diagnosis are achieved by propagating the information in forward and backward directions, which means the causal network can not only calculate the pipeline segment risk given the context conditions, but also recommend the optimal mitigation approaches taken to achieve certain operation goals. The network developed in this section acts as the engine for the enterprise decision support system. The detailed description of the overall development process is discussed below.

Bayesian Network Theory

Bayesian networks are causal probabilistic models that combine data and subject matter expert knowledge to quantify certainty, providing the most rigorous and rational basis for critical decision making process. Specifically, a Bayesian network is a directed acyclic graph where the nodes represent a set of random variables $X = \{X_1, X_2, X_3, \dots, X_n\}$, and the edges connecting any two nodes represent direct dependency between them[19]. The strength of an edge is given by conditional probability distribution of the nodes associated with the edge. **Figure 4-1** shows a Bayesian network model of five nodes.

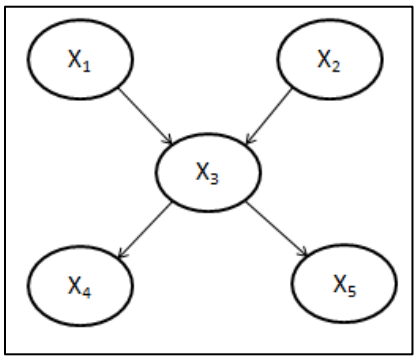


Figure 4-1. Example of Bayesian network with five nodes

Mathematically, a Bayesian network of a set of n random variables $X = \{X_1, X_2, X_3, \dots, X_n\}$ is represented by its joint probability distribution $P(X_1 = x_1, X_2 = x_2, X_3 = x_3, \dots, X_n = x_n)$ or $P(x_1, x_2, x_3, \dots, x_n)$, where $x_1, x_2, x_3, \dots, x_n$ are the values of the variables $X_1, X_2, X_3, \dots, X_n$,

respectively. The probability distribution for each variable X_i is computed by taking the marginal integration of all the other variables. Based on the direction of information propagation, the Bayesian network can be used for prognosis and diagnosis.

$$\begin{aligned} &P(x_1, x_2, x_3, \dots, x_n) \\ &= P(x_1)P(x_2|x_1)P(x_3|x_1, x_2) \dots P(x_n|x_1, x_2, x_3, \dots, x_{n-1}) \\ &= \prod_{i=1}^{n} P(x_i|x_1, x_2, x_3, \dots, x_{i-1}) \end{aligned}$$

Equation 4-1

Pipeline Risk Modeling

It has been stated before that the failure of pipeline results from the interaction of multiple causes. The capabilities of Bayesian network enable the calculation of threat interaction levels and severity of risks in vintage gas pipelines. To improve the analysis accuracy, the subject matter expert's opinions, historic data, and concurrent inspection data are integrated in the developed Bayesian network model.

Graphical Analysis for Risk Factors

Initially, 32 factors, identified by SMEs were considered to have critical impact on the performance of vintage gas pipelines. Those factors are illustrated in **Figure 4-2** and they represent the nodes of the developed Bayesian network.

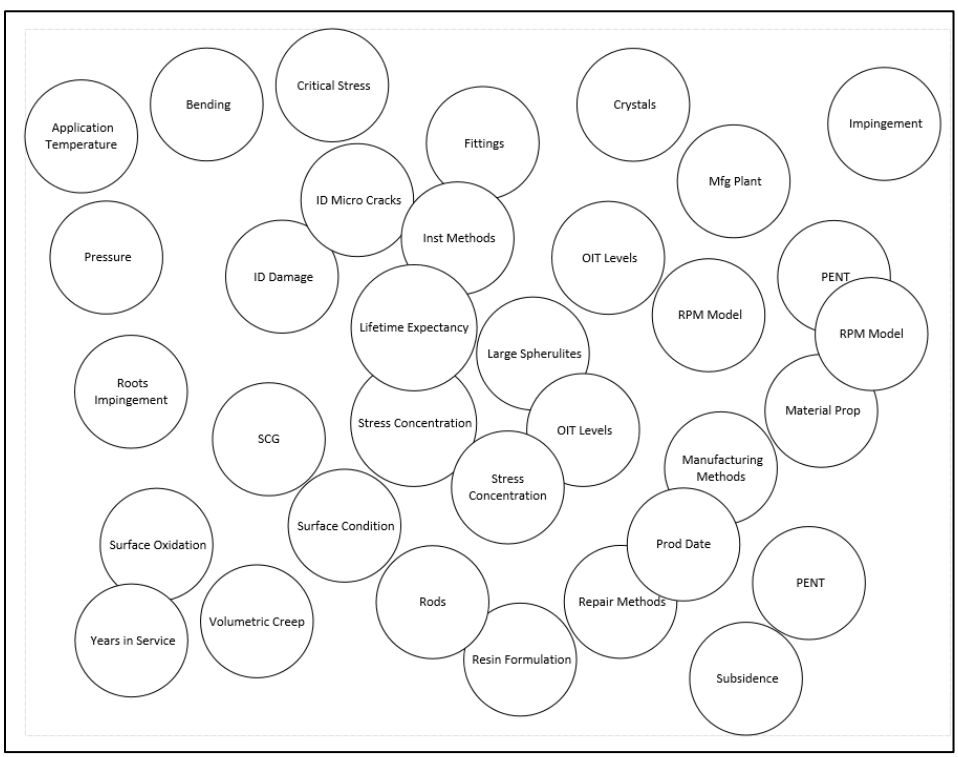


Figure 4-2. Factors selected by SMEs

The second step is to determine the interaction between every two nodes. Out of a theoretical possible combination of 4,294,967,296 interactions between the 32 factors, SMEs then selected 65 plausible interactions to construct the edges of the initial Bayesian network as shown in **Figure 4-3**. Two simple graphical analyses were then run. First, betweenness centrality, which is equal to the number of shortest paths from all vertices to all others that pass through that node, was calculated for every node. It is shown in **Figure 4-4**. Second, degree centrality, which is the number of ties a node has to other nodes, was calculated for each node as shown in **Figure 4-5**. Larger red dot denotes higher centrality value in both **Figure 4-4** and **Figure 4-5**. The product of degree centrality and betweenness centrality gives a composite ranking, which then helps to identify the most important factors. **Figure 4-6** and **Figure 4-7** show the composite ranking of nodes in the Bayesian network. As expected, the top five factors with the highest composite rankings in decreasing order are Stress concentration, Manufacturing methods, Bending, RPM model, and Surface condition.

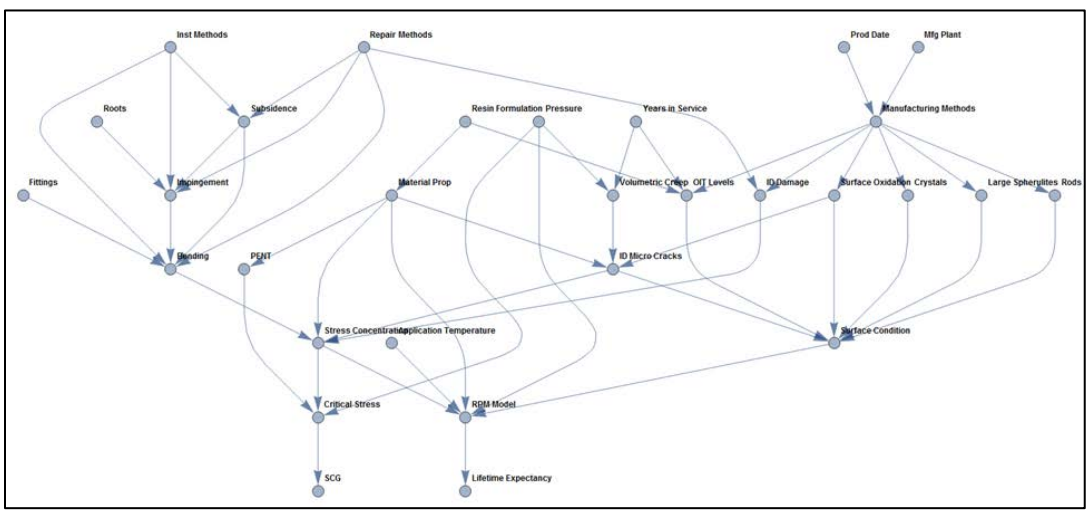


Figure 4-3. Initial Bayesian network graph selected by SMEs

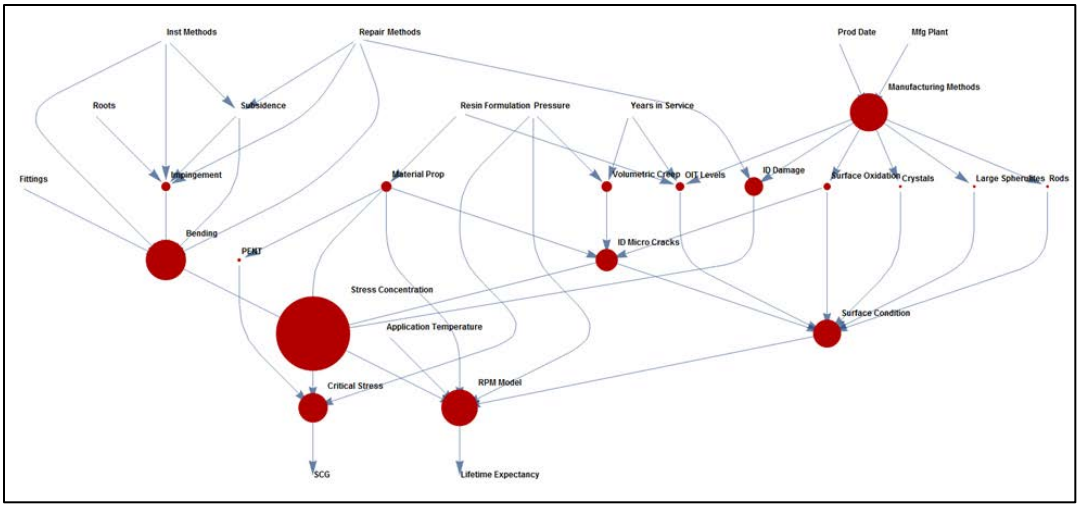


Figure 4-4. Betweenness centrality of the Bayesian network graph

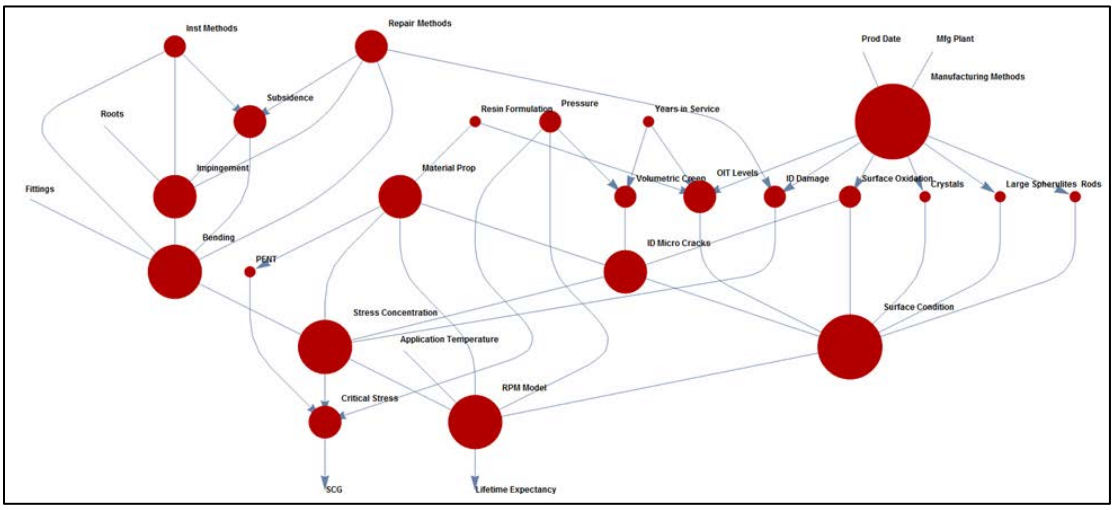


Figure 4-5. Degree centrality of the Bayesian network graph

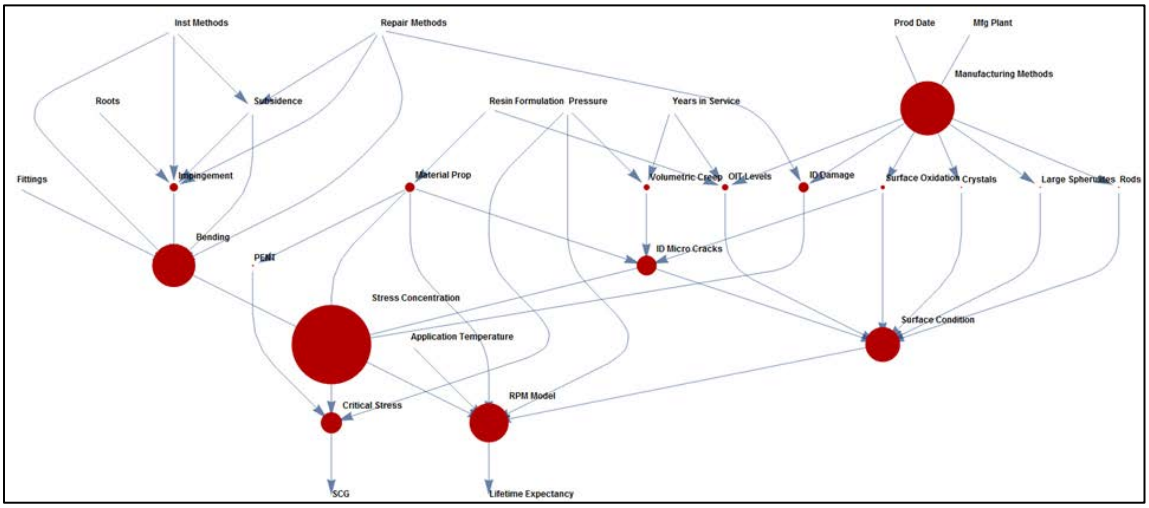


Figure 4-6. Composite ranking of the Bayesian network graph

Stress Concentration	51.
Bending	27.5
Manufacturing Methods	26.
RPM Model	25.
Critical Stress	20.
Surface Condition	19.
ID Micro Cracks	15.
ID Damage	12.5
Volumetric Creep	7.
Material Prop	7.
Impingement	6.
OIT Levels	5.6
Surface Oxidation	4.6
PENT	2.
Rods	1.6
Large Spherulites	1.6
Crystals	1.6
Years in Service	0.
Subsidence	0.
SCG	0.
Roots	0.
Resin Formulation	0.
Repair Methods	0.
Prod Date	0.
Pressure	0.
Mfg Plant	0.
Lifetime Expectancy	0.
Inst Methods	0.
Fittings	0.
Application Temperature	0.

Betweenness Centrality

Manufacturing Methods	8
Surface Condition	7
Stress Concentration	6
RPM Model	6
Bending	6
Material Prop	5
Impingement	5
ID Micro Cracks	5
Subsidence	4
Repair Methods	4
OIT Levels	4
Critical Stress	4
Volumetric Creep	3
Surface Oxidation	3
Pressure	3
Inst Methods	3
ID Damage	3
Years in Service	2
Rods	2
Resin Formulation	2
PENT	2
Large Spherulites	2
Crystals	2
SCG	1
Roots	1
Prod Date	1
Mfg Plant	1
Lifetime Expectancy	1
Fittings	1
Application Temperature	1

Degree Centrality

Stress Concentration	306.
Manufacturing Methods	208.
Bending	165.
RPM Model	150.
Surface Condition	133.
Critical Stress	80.
ID Micro Cracks	75.
ID Damage	37.5
Material Prop	35.
Impingement	30.
OIT Levels	22.4
Volumetric Creep	21.
Surface Oxidation	13.8
PENT	4.
Rods	3.2
Large Spherulites	3.2
Crystals	3.2
Years in Service	0.
Subsidence	0.
SCG	0.
Roots	0.
Resin Formulation	0.
Repair Methods	0.
Prod Date	0.
Pressure	0.
Mfg Plant	0.
Lifetime Expectancy	0.
Inst Methods	0.
Fittings	0.
Application Temperature	0.

Composite Ranking = Betweenness Centrality * Degree Centrality

Figure 4-7. Composite ranking of the nodes in Bayesian network graph

The third step is to construct conditional probability tables for the factors in the Bayesian network. Three methods were identified for the purpose: FEM analysis, Data collection and analysis, and Experimentation and observation as shown in **Figure 4-8**. We used an extensive historic hydrostatic test data covering all vintage pipelines to construct the preliminary conditional probability table. **Figure 4-9** demonstrates the preliminary analysis results. The initial study was conducted on a set of 105 slow crack growth failures in pipe that were not associated with fittings or other known stress risers. The pipes were all extracted from a gas distribution system in 2010 -2011. They had been in service since 1972, 1973 and 1974. A combination of simpler statistical models, subject matter expertise and historical observations was used to estimate initial graphical structure and probabilistic distribution of a Bayesian network model. This process is represented in ontology illustrated by **Figure 4-10**. Statistical models help to understand various levels of threat interactions between factors of interest that may lead to failure. Some of the modeling methods, evaluated in this preliminary work, are rate process method, fault tree analysis, event tree analysis, surface feature model, finite element method, risk score model, etc. The following sections will demonstrate the continuous refinement of the developed Bayesian network when additional data or knowledge are collected.

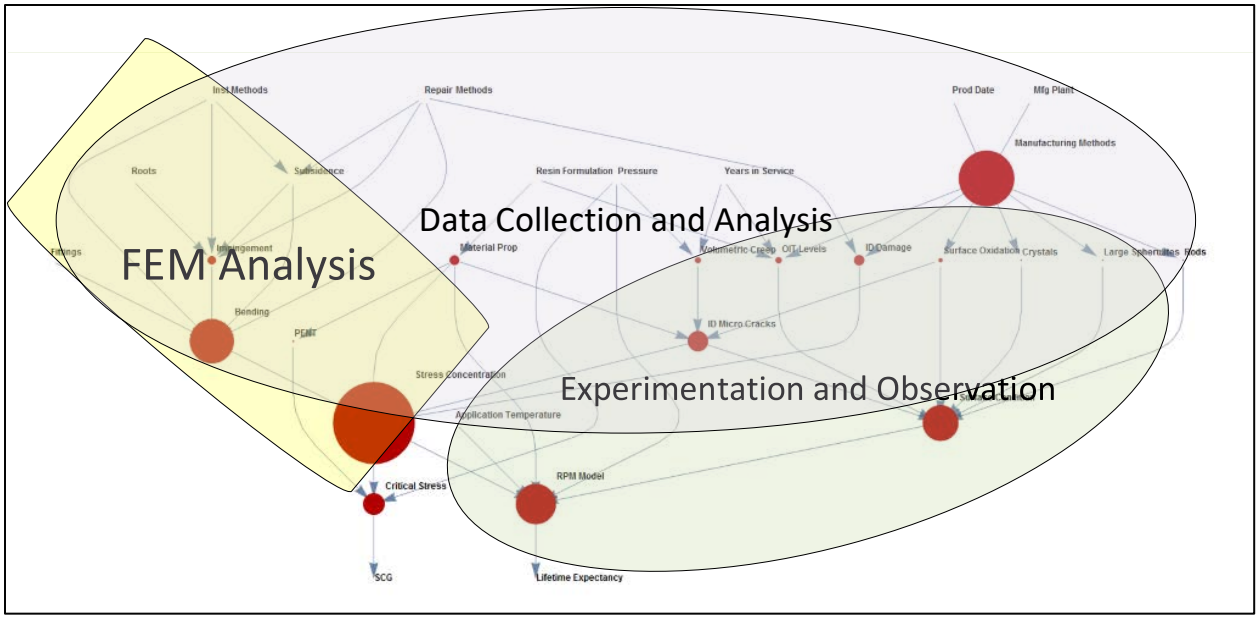


Figure 4-8. Methods to derive the conditional probability tables in the Bayesian network

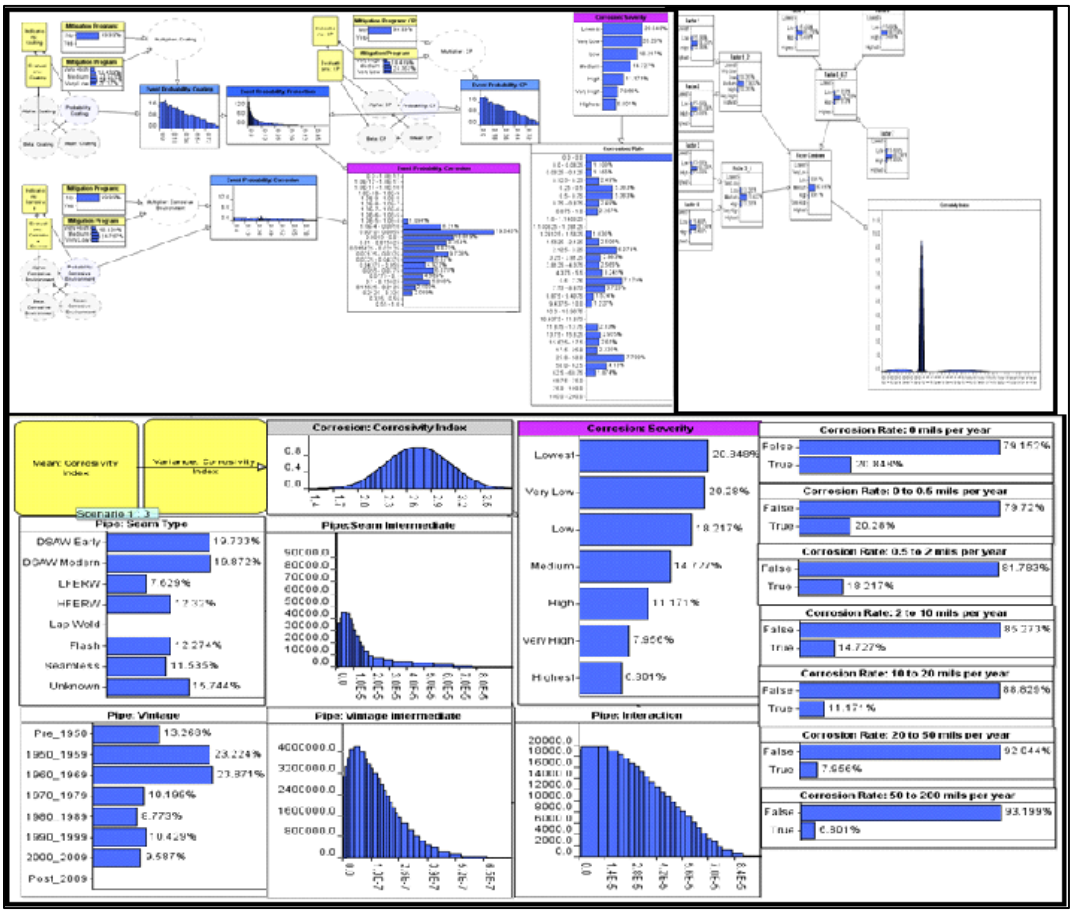


Figure 4-9. Preliminary conditional probability distributions

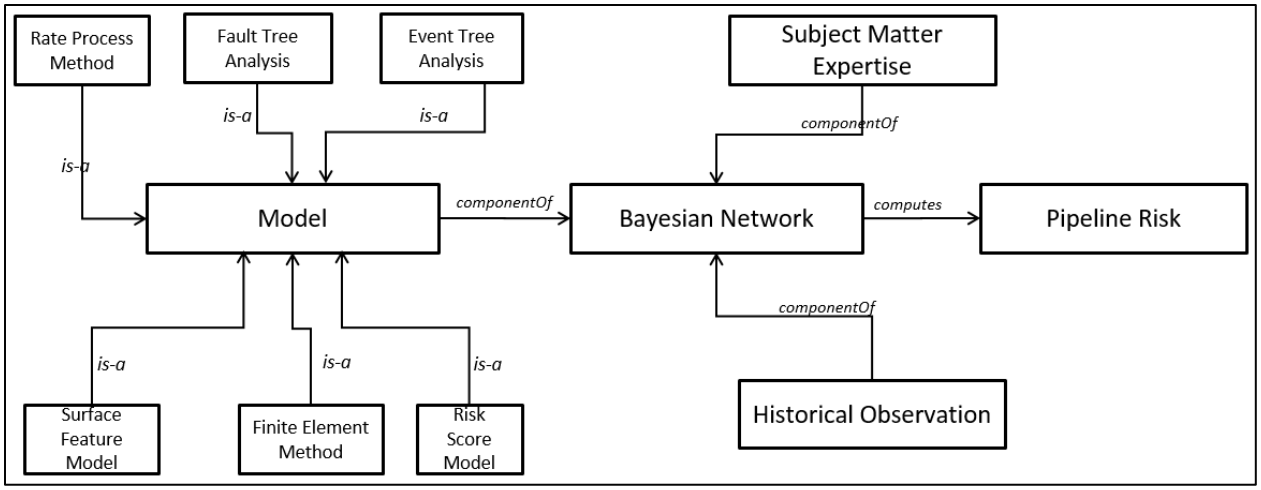


Figure 4-10. Ontology describing the models, subject matter expert's knowledge, historical observation, and Bayesian network

Bayesian Network Quantification for Aldyl A Risk Assessment

Under the normal operation pressure, pipelines will experience the slow crack growth (SCG) failure pattern, because the applied stress is too low to cause ductile failure. Three factors were identified to have the greatest impact on the pipe SCG lifetime expectancy - pipe inner wall surface condition, stress state, and ground temperature. The pipe inner surface condition depends on a number of factors notably the presence of surface oxidation, surface crystals, boundary crystals, large Spherulites, rods, inside diameter micro cracks (IDMC), and dimple. The pipe stress state is the result of any stress risers acting on the pipe such as material properties, impingement, bending, etc. Ground temperature is directly associated with the creep behavior of a pipe, and thus affects its lifetime expectancy.

In this project, AgenaRisk[20], a popular risk analysis and decision support software, is used to develop the causal network, which is then deployed to the developed Enterprise Decision Support System. The JavaScript Object Notation (JSON) messaging format from the previous quarter is also updated with new features for sensitivity analysis, initialization, update and query operation in the Bayesian network model. Given the SCG failure pattern described above, the initially constructed Bayesian network in **Figure 4-3** is updated and illustrated in **Figure 4-11**. In this updated network, volumetric creep node is eliminated because through extensive studies, available volumetric creep models could not provide satisfactory prediction for the initiation of inner diameter micro cracks (IDMC) as well as the total lifetime expectancy. An alternative approach is proposed in the following section to incorporate the effect of years in service. Apart from that, some intermediate nodes, such as "OD_ID_ratio", "Resultant Movement", and "Risk Score" are introduced here to alleviate computational efforts.

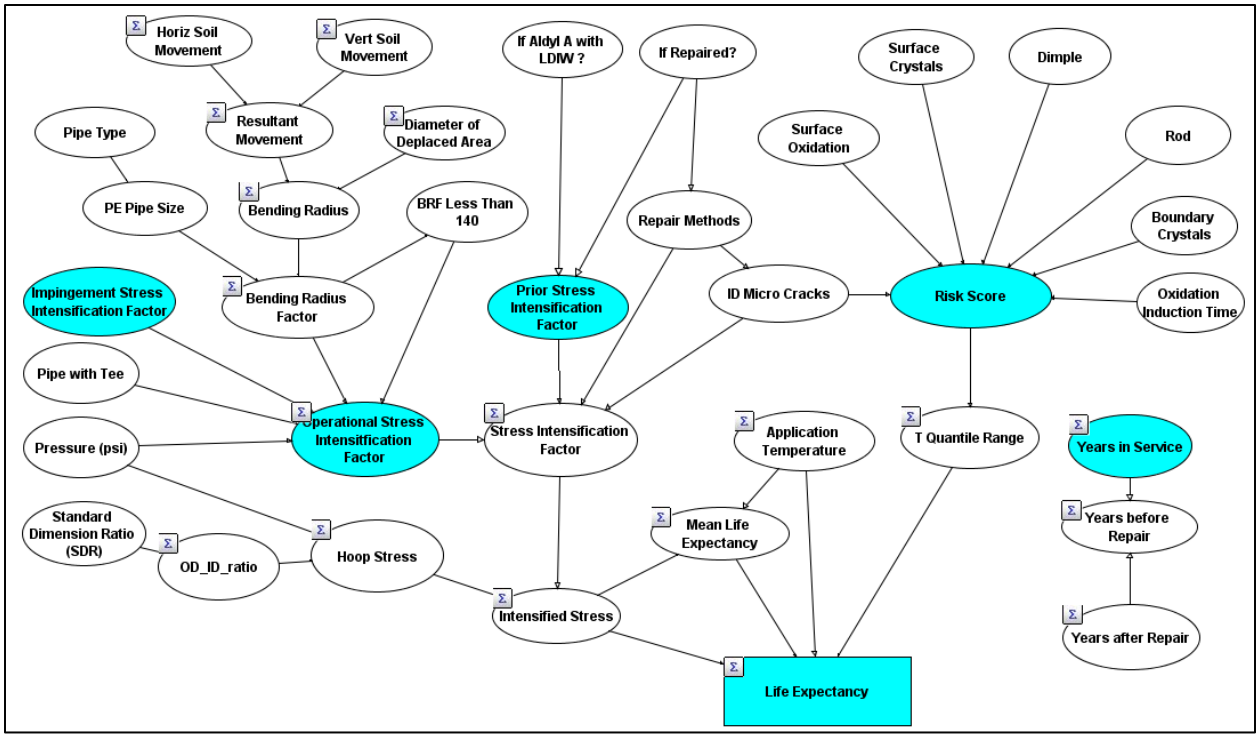


Figure 4-11. The modified overall Bayesian network

Prior Stress Intensification Factor (PSIF)

In this section, the effect of years in service is considered by introducing the prior stress intensification factor (PSIF) node. The stress concentration node in the initial design is modified as operational stress intensification factor (OSIF), which represents the stress intensification factor (SIF) caused by factors related to pipeline operations, such as soil movement, impingement, pipe type, operational pressure, etc. Analogous to previous discussion, the network is partitioned into several sub-networks according to the nodes highlighted in **Figure 4-11**. The node "Impingement Stress Intensification Factor" is the output from a separate Bayesian network model, which will be discussed in detail below. The node types and states are determined with the best knowledge of the investigator and they are summarized in **Table 4-1**.

Table 4-1. The summary for node types and states

Node Name	Node Type	States/Distribution	Unit	Explanation
Years in Service	Continuous	N (40, 400, 5, 70)*	Year	From Installation Time to Now
Years after Repair	Continuous	U (0, 50)**	Year	From Repair Time to Now
Years before Repair	Continuous		Year	Years-YaftRep
If Aldyl A with LDIW	Boolean	(False, True)		Check If the Aldyl a Pipe Has Low Ductile Inner Wall (LDIW)

Node Name	Node Type	States/Distribution	Unit	Explanation
If Repaired	Boolean	(False, True)		Check If the Pipe Has Been Repaired
Repair Methods	Labelled	(Mechanical Coupling, Butt Fusion, NA)		Repair Methods
Prior Stress Intensification Factor (PSIF)	Continuous			Prior Stress Intensification Factor based on Historic AldylA Data
Horiz Soil Movement	Continuous	N(12, 36, 0, 120)	Ft	Soil Movement along Horizontal Direction
Vert Soil Movement	Continuous	N(24, 36, 0, 120)	Ft	Soil Movement along Vertical Direction
Resultant Movement	Continuous		Ft	Resultant Soil Movement
Diameter of Displace Area	Continuous	N(50,100, 0, 100)	Ft	Diameter of the Area with Soil Movement
Bending Radius	Continuous			Pipe Bending Radius
Pipe Type	Boolean	(Service Lines, Mains)		Pipe Function Type
PE Pipe Size	Labelled	(<=1 inch, <=2inch, <=4 inch, <= 6 inch)		Pipe Outside Diameter
Bending Radius Factor	Continuous			Bending Radius/Pipe Size
BRF Less Than 140	Boolean	(False, True)		Check if the Bending Radius Factor is Less Than 140
Pipe with Tee (PWT)	Boolean	(False, True)		Check if the AldylA Pipe has a Tee
Operational Stress Intensification Factor (OSIF)	Continuous			SIF Caused by Root, Rock, Impingement, and Bending
Stress Intensity Factor (SIF)	Continuous			Combination Between PSIF and OSIF
ID Micro Cracks	Discrete Real	(0.0, 1.0)		Existence of Micro Cracks on Inside Diameter
Surface Oxidation	Discrete Real	(0.0, 1.0, 2.0, 3.0, 4.0, 5.0)		Surface Oxidation on Inside Diameter
Surface Crystal	Discrete Real	(0.0, 1.0, 2.0, 3.0, 4.0, 5.0)		Surface Crystal on inside diameter
Dimple	Discrete Real	(0.0, 1.0)		Existence of Dimple on Inside Diameter
Rod	Discrete Real	(0.0, 1.0)		Existence of Rod on Inside Diameter
Boundary Crystal	Discrete Real	(0.0, 1.0, 2.0, 3.0, 4.0, 5.0)		Boundary Crystal on Inside diameter
Risk Score	Labelled	(1, 2, 3, 4)		Relative Risk Score Determined by Pipe Micro Surface Features
T Quantile Range	Continuous	([-1.29,6.50],[-2.36,-1.29], [-3.84,-2.36],[-5.00,-3.84])		Quantile Range from t Distribution
Application Temperature	Continuous	23	Celsius	Average Operation Temperature of the Pipe
Pressure	Continuous	N(50, 400, 1, 100)	Psi	Operation Pressure
Standard Dimension Ratio (SDR)	Discrete Real	(9, 11, 13)		Outside Diameter/Wall Thickness

Node Name	Node Type	States/Distribution	Unit	Explanation
OD_ID_Ratio	Continuous	(1.6531, 1.4938, 1.3967)		SDR^2/(SDR-2)^2
Hoop Stress	Continuous		Psi	Hoop Stress Caused by Internal Pressure
Intensified Stress	Continuous		Psi	Hoop Stress Intensified by SIF
Mean Life Expectancy	Continuous		Log(hrs)	Mean Life Prediction using Control SCG Model
Life Expectancy	Continuous		Log(hrs)	Life Prediction Considering Risk Score of the Pipe
Impingement SIF	Continuous			SIF Caused by Impingement
* Represents truncated normal distribution $N(\mu, \sigma^2, Lower\ Bound, Upper\ Bound)$ **Represents uniform distribution $U(Lower\ Bound, Upper\ Bound)$				

The Dupont Company has performed hydrostatic testing of Aldyl A in accordance with the industry accepted Rate Process Model (RPM). The times of failure of the specimens under different conditions overlay the Dupont control reference data at all stress levels. The hydrostatic pressure testing data sets for Aldyl A are depicted in **Figure 4-12**, in which it can be seen that the pipes have different failure time distributions for different geometry/loading conditions such as low ductile inner wall (LDIW), indentation/impingement, and squeeze-off. LDIW is the inner surface feature defects induced by manufacturing process for Aldyl A in 1970s. Since indentation has similar effect with the impingement, its effect is considered in the impingement node. Generally, squeeze-off is a common method for stopping the gas flow for plastic pipe repair. Therefore, it is assumed that the pipe will be under squeeze-off loading condition if the pipe has been repaired.

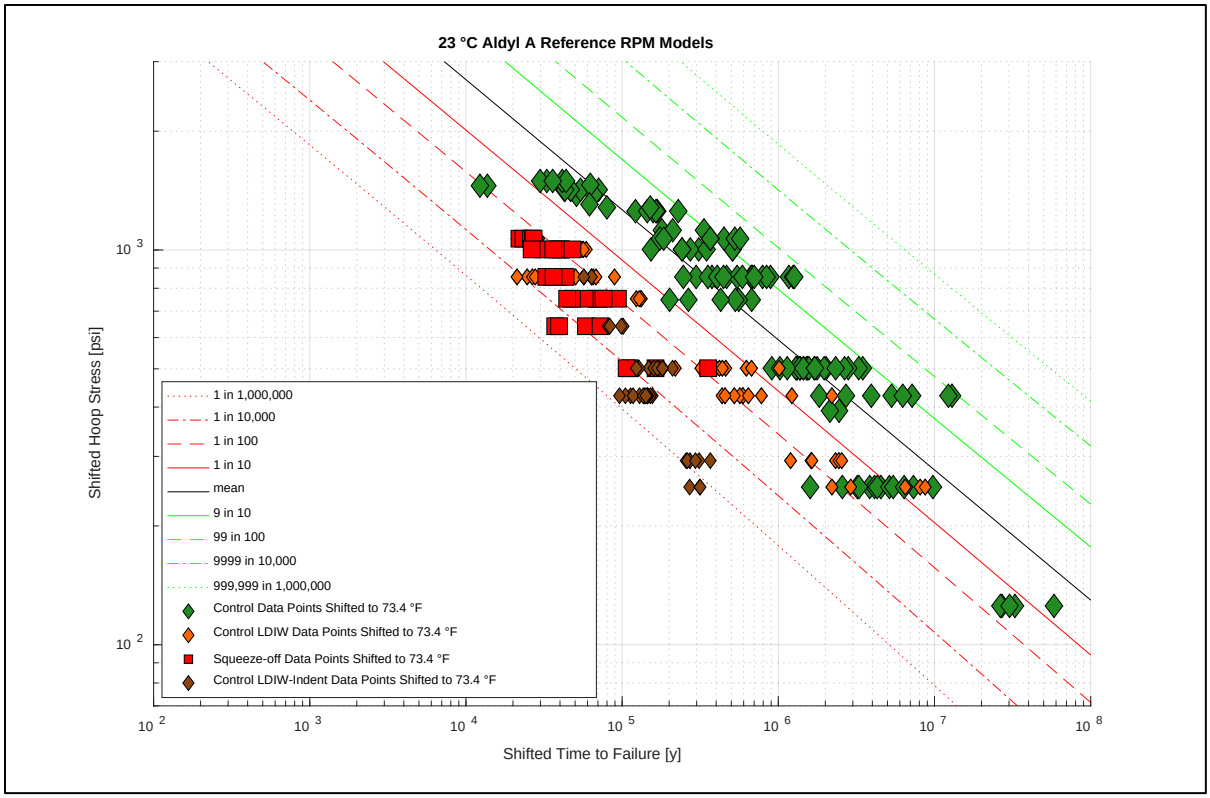


Figure 4-12. Data sets for SCG control, LDIW control, LDIW with indentation and LDIW with squeeze-off shifted to 73.4°F

In pressure testing guidelines, stress intensification factor (SIF) originating from the extrusion process seeds the ductile failure mode, which was consistent with the results from finite element analysis (FEA). Bearing the idea that the distribution of SIF on the inside diameter will not vary between different testing methods, the mean SIF value for the SCG data should be identical with that of the Ductile data. To compute the life under different loading conditions, the data sets for control LDIW, LDIW with Indentation, and LDIW with squeeze-off were mapped to the SCG control curve with an equivalent SIF value for each data point. The mapping process is discussed in **8** below. The prior distributions for SIF under different loading conditions are depicted in **Figure 4-13**. Before substituting the computed SIF distribution into the RPM, the values need to be normalized with respect to its mean SIF value of the control SCG data and the normalized SIF is represented as SIFN below.

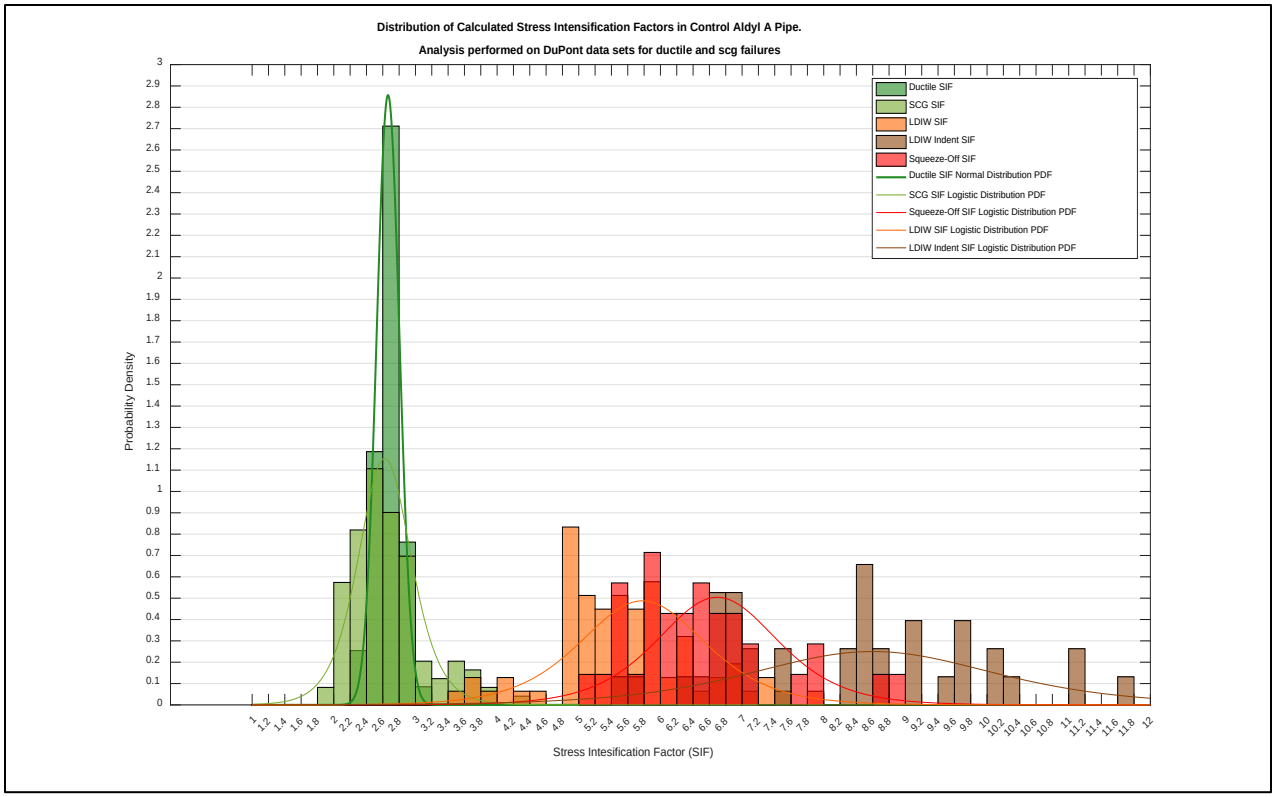


Figure 4-13. The SIF distributions under different loading conditions

To accurately quantify the remaining useful life (RUL) of pipes, it is critical to consider the effect of years in service. The schematics for computing the relative stress intensification factor (SIFR) are illustrated in **Figure 4-14**. Plugging these two data points into SCG reference rate process model (RPM), we can get

$$\log(t_0) = C_1 + \frac{C_2}{T} + \frac{C_4}{T}\log(\sigma_0) \qquad \textbf{Equation 4-2}$$

$$\log(t_0 - \Delta t) = C_1 + \frac{C_2}{T} + \frac{C_4}{T}\log(\sigma_1) \qquad \textbf{Equation 4-3}$$

Combining Equation **4-2** and **Equation 4-3**, the relative SIF can be computes as

$$SIFR = \frac{\sigma_1}{\sigma_0} = \left(\frac{t_0 - \Delta t}{t_0}\right)^{\frac{T}{C_4}} \qquad \textbf{Equation 4-4}$$

Thus, the total SIF (SIFT) after Δt years in service is expressed as

$$SIFT = SIFR \times SIFN \qquad \text{Equation 4-5}$$

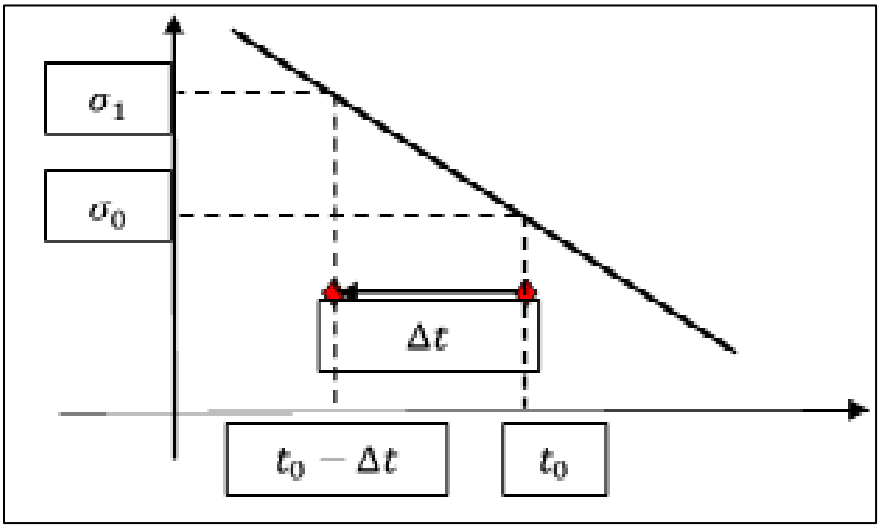


Figure 4-14. The schematics for computing the relative SIF (SIFR)

Given the years in service, the total failure life of the pipe is the summation of years in service and the predicted RUL. Taking the LDIW data as an example, **Figure 4-15** illustrates the procedures for computing the RUL and total failure time after certain years in service. The LDIW is shifted to the left by Δt years as shown in **Figure 4-16**. Conceptually, the predicted total failure time after certain years in service should be identical to the control LDIW data without considering other material degradation factors. The predicted total failure time considering different years in service is given in **Figure 4-17**.

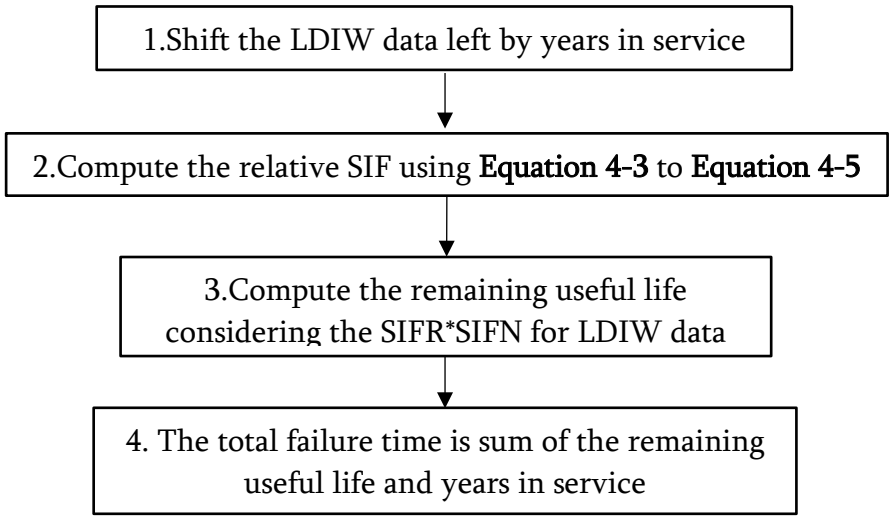


Figure 4-15. The procedures for computing the RUL and total failure time

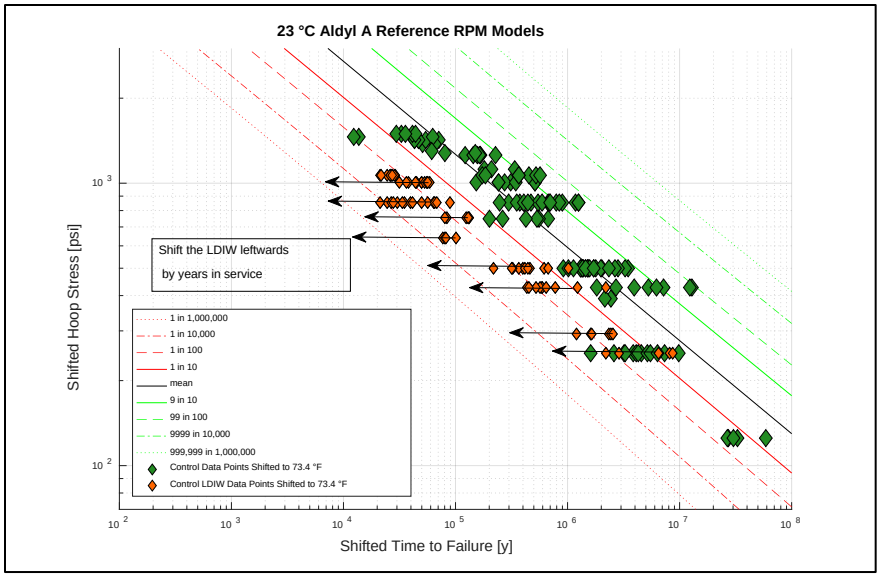


Figure 4-16. Shift the LDIW data leftwards by Δt

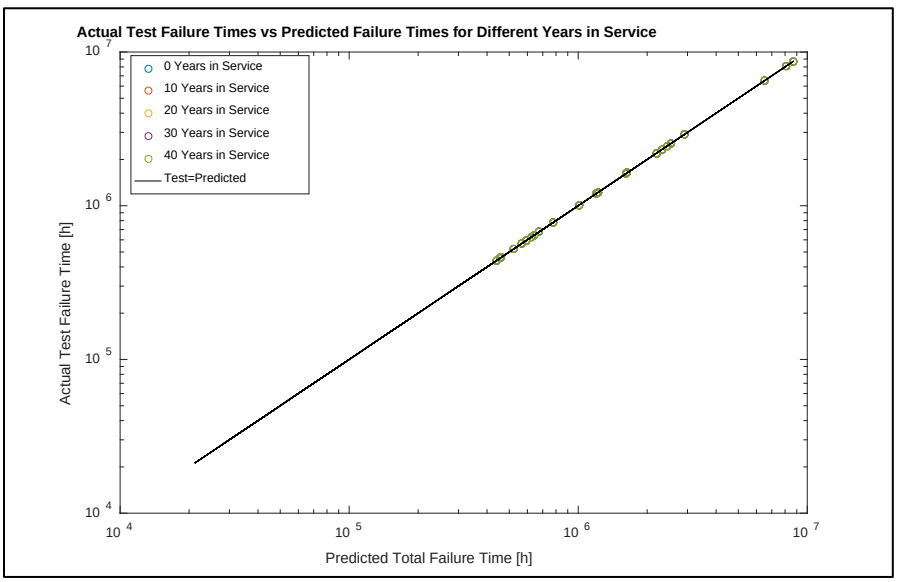


Figure 4-17. The validation of predicted total failure time after different years in service

For pipes in service, their loading conditions may change due to the change of pipe inner wall surface features, such as LDIW or internal damage caused by squeeze-off. In the developed Bayesian network, different loading conditions generated by different combinations of repair or inner diameter surface features are summarized in **Table 4-2**. That means different data sets will be used for computing the SIFT in **Equation 4-5** for each individual case. The SIFT are proportionally combined by the ratio of the years under a specific loading condition and the total years in service. For example, the pipe has been in service for Δt years under the loading condition 1 and loading condition 2 for Δt_1 and Δt_2 years, respectively. The final SIFT can be calculated as

$$SIFT = SIFT_1 \frac{\Delta t_1}{\Delta t} + SIFT_2 \frac{\Delta t_2}{\Delta t}$$ **Equation 4-6**

The sub-network for computing the prior SIF distribution is shown in **Table 4-2**, in which the prior SIF distribution is computed considering the years in service as well as the years after repair by means of AgenaRisk application programming interface (API)

Table 4-2. Different combinations of loading conditions

If LDIW? / If Repaired?	True	False
True	LDIW + Squeeze off	Control SCG + Squeeze off
False	LIDW	Control SCG

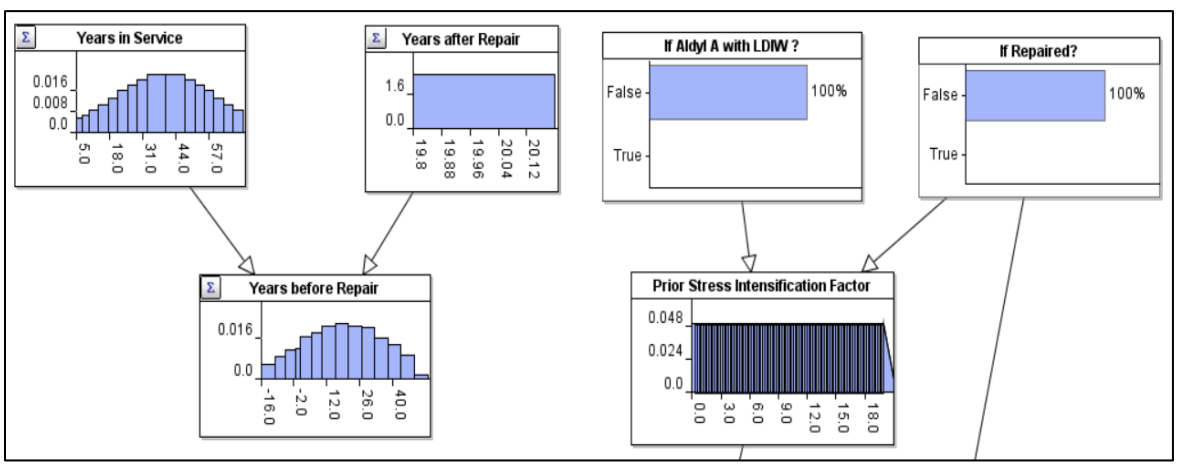


Figure 4-18. The sub-network for computing the prior SIF distribution

Impingement Stress Intensification Factor (ImpSIF)

Since the deformation caused by indentation and impingement are similar, they are generally considered as impingement in this developed Bayesian network. In 1982, Dupont issued a letter urging its customers to realize the risk of pipes subject to rock/root impingement. According the report published in 2014[21], the stress concentration introduced by impingement occurs more often than that caused by squeeze-off. Additionally, impingement is of particular concern because it would accelerate pipe failures by around 12 years from the time when the pipe was initially impinged.

Extensive studies have been performed to quantify the stress or strain caused by indentation/impingement[22, 23]. In[24], the author applied the equivalent load approach to study the stress distribution around imperfection in thin-walled structures. The latest version of ASME B31.8 standard recommends the equations for calculating the effective strain in

dents. Based on the theory of plates, the strain components are estimated considering the geometry of the deformation, which includes the dent length and dent depth along the circumferential and axial directions.

$$\varepsilon_1 = \frac{t}{2}\left(\frac{1}{R_o} - \frac{1}{R_1}\right) \qquad \textbf{Equation 4-7}$$

$$\varepsilon_2 = -\frac{1}{2}\frac{t}{R_2} \qquad \textbf{Equation 4-8}$$

$$\varepsilon_3 = \frac{1}{2}\left(\frac{d}{L}\right)^2 \qquad \textbf{Equation 4-9}$$

where ε_1, ε_2, ε_3 stand for the circumferential bending strain, longitudinal bending strain, and extensional membrane strain. The values for R_1 and R_2 are negative for reentrant impingement. The symbols in above equations illustrated by plotting them in **Figure 4-19** below,

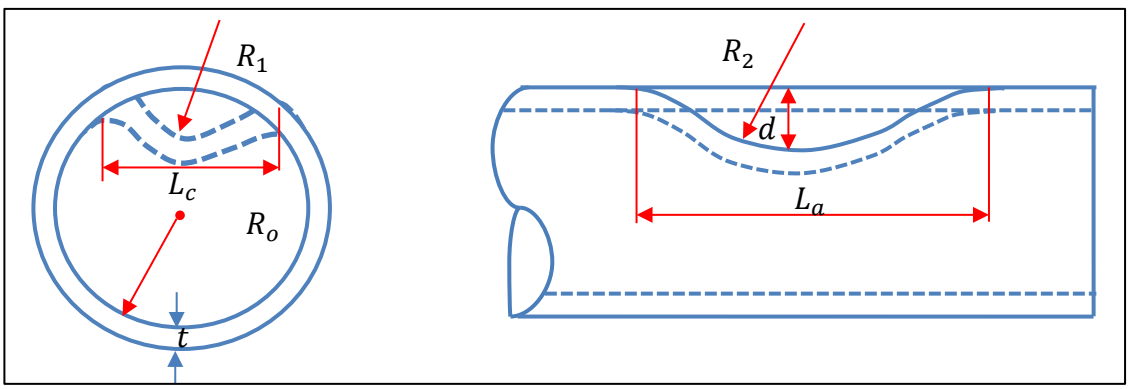


Figure 4-19. The geometry of indentation/impingement

As stated in[25], the pipe material will experience severe plastic deformation due to indentation, therefore the elasto-plastic characteristic of the polyethylene material should be used to avoid overestimating the stress concentration. GTI has generated the stress strain curves for medium density polyethylene and Aldyl A materials, which was successfully implemented to simulate the nonlinearity of the squeeze-off process. The stress strain curve for Aldyl A material is illustrated in **Figure 4-20**.

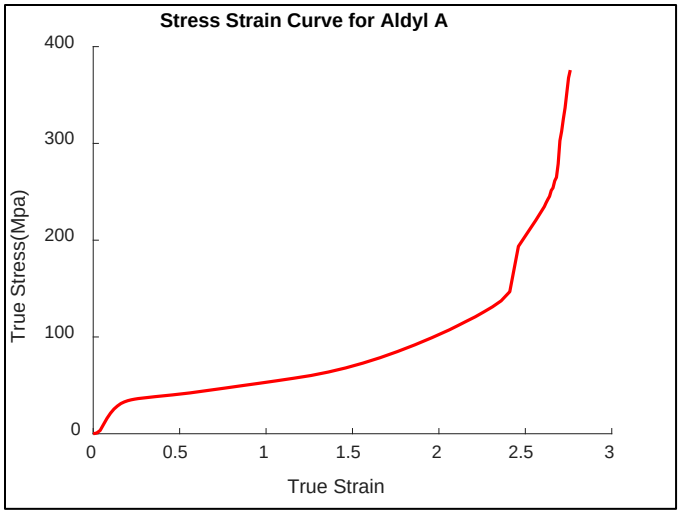


Figure 4-20. The stress strain curve for Aldyl A

Since the pipe is not undergoing uniaxial stress state, the strain components along tri-axial directions are combined to calculate the effective strain in the inside diameter,

$$\varepsilon_{eff} = \frac{\sqrt{2}}{3}\sqrt{(e_1 - e_2)^2 + (e_1 - e_3)^2 + (e_2 - e_3)^2} \qquad \textbf{Equation 4-10}$$

where e_1, e_2, e_3 stand for the total strain along circumferential, axial, and radial directions, respectively and ε_{eff} is the combined effective strain. For polyethylene, the Poisson's ratio is very close to 0.5, Equation 4-10 can be simplified as $\varepsilon_{eff} = e_1$ under uniaxial loading. Generally, the stress along radial direction can be neglected in comparison with the other two directions and Equation 4-10 is rearranged as

$$\varepsilon_{eff} = \frac{2}{3}\sqrt{{e_1}^2 - e_1 e_2 + {e_2}^2} \qquad \textbf{Equation 4-11}$$

where e_1=ε_1 and $e_2 = \varepsilon_2 + \varepsilon_3$. Using the effective strain, the effective stress induced by impingement can be computed by substituting the effective strain in the stress strain curve.

Empirically, for a specific pipe diameter, the geometrical parameters shown in **Figure 4-19** are bounded. For example, the dent depth d should be less than the pipe radius R_o. It is defined in that dents are considered to be injurious if the depth d is greater than 6%. The circumferential dent length L_c should not be greater than the pipe outside diameter. Based on the geometric relationship, the bend radius along the circumferential and axial directions can be estimated as

$$R_1 = \frac{L_c^2}{8d} + \frac{d}{2} \qquad \textbf{Equation 4-12}$$

$$R_2 = \frac{L_a^2}{8d} + \frac{d}{2}$$ **Equation 4-13**

Thus, the sub-network for computing strain components is shown in **Figure 4-21**. It is explicitly demonstrated that the circumferential strain is much higher in comparison with the other two components, which can be explained by the restrained dent length along the circumferential direction. If impingement information is unknown to the operator, the induced SIF severity can be empirically estimated using linear combination of root and rock density. Given the prior distributions, the distribution of impingement induced stress concentration can be calculated and shown in **Figure 4-22**(a). **Figure 4-22**(b) plots the SIF distribution by processing the reference Aldyl A hydrostatic testing data.

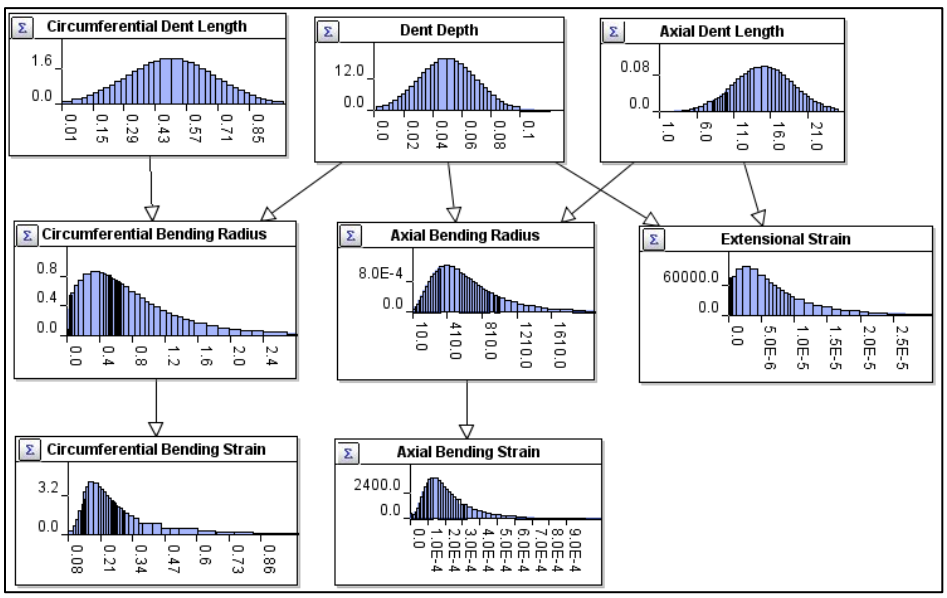
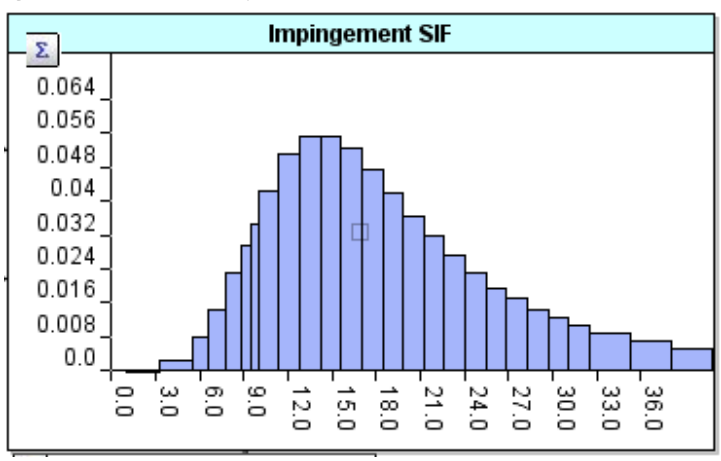
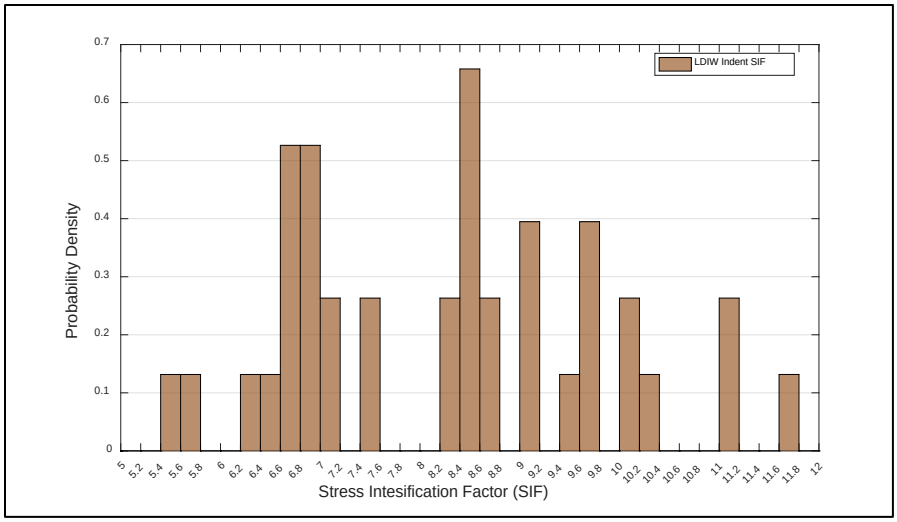


Figure 4-21. The Bayesian network for strain components



(a)



(b)

Figure 4-22. The distributions of SIF from Bayesian network (a) and experiment (b)

From the above plots, it can be observed that the estimation from the Bayesian network is much higher than the experimental results. The discrepancy may originate from the accuracy for the bending radius estimation and the true stress strain curve difference for the pipe under hydrostatic testing. For simplicity, a correction factor 0.5 is introduced here in the Bayesian network to make it consistent with the experimental testing. After calibration, the Bayesian network for computing the impingement SIF is depicted in **Figure 4-23**.

Table 4-3 summarizes the type and states for each node.

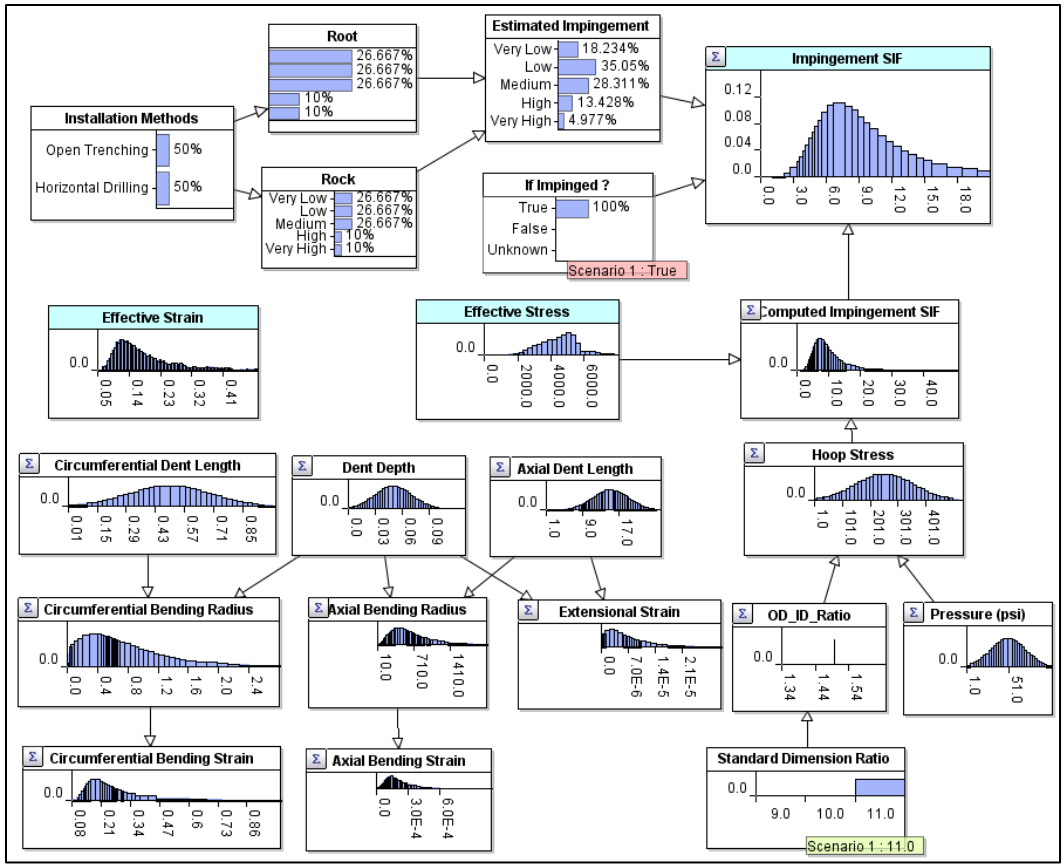


Figure 4-23. The Bayesian network model for computing impingement SIF

Table 4-3. The type and states for each node

Node Name	Node Type	States/Distribution	Unit	Explanation
Circumferential Dent Length	Continuous	N(0.5, 0.04, 0, 1)	In	Dent Length along Circumferential Direction
Dent Depth	Continuous	N(0.05, 0.0004, 0.00001, 0.14)	In	Depth of the Dent
Axial Dent Length	Continuous	N(15, 16, 0.5, 25)	In	Dent Length along Axial Direction
Circumferential Bending Radius	Continuous		In	$\lvert R_1 \rvert$
Axial Bending Radius	Continuous		In	$\lvert R_2 \rvert$
Circumferential Bend Strain	Continuous			ε_1
Axial Bending Strain	Continuous			ε_2
Extensional Strain	Continuous			ε_3
Effective Strain	Continuous			ε_{eff}
Effective Stress	Continuous		Psi	Stress on Corresponding to ε_{eff} on the Stress Strain Curve

Node Name	Node Type	States/Distribution	Unit	Explanation
Computed Impingement SIF	Continuous			Effective Stress/Hoop Stress
Installation Methods	Labelled	(Open Trenching, Horizontal Drilling)		The Methods of Installation
Root	Ranked	(Very Low, Low, Medium, High, Very High)		Root Density
Rock	Ranked	(Very Low, Low, Medium, High, Very High)		Rock Density
Estimated Impingement	Ranked	(Very Low, Low, Medium, High, Very High)		Estimated Impingement Severity
If Impinged ?	Labelled	(Yes, No, Unknown)		Check if the Pipe is Impinged
Impingement SIF	Continuous			SIF Caused by Impingement

As seen above, the distribution for "Impingement SIF" node overlays the distribution from experimental testing after introducing the correction factor. The advantage of this network is it can provide relatively accurate SIF prediction for different impingement configurations, which is more informative than the extracted SIF distribution for the decision making process.

Operational Stress Intensification Factor (OSIF)

Operational stress intensification factor (OSIF) represents the SIF introduced by real operational conditions, such as impingement, soil movement or bending. As stated previously, prior stress intensification factor (PSIF) represents the SIF introduced by primary loading conditions caused by years in service or repair methods. The superposition of these two sources of SIF is the total stress intensification factor (SIF).

A FEA model was created to investigate the SIF caused by bending, with or without fitting tees. Then, the model was extended to consider the relative angle between fitting tees and bending curvature. **Figure 4-24** shows the definition of bending sign convention used in this report. The tee will be on the concave up portion of the pipe for positive bending and on the convex up portion of the pipe for negative bending. For each pipe configuration, the SIF was calculated by taking the maximum von Mises stress in each load case and dividing it by the nominal pipe hoop stress calculated from the internal pressure of the respective load case. After that, a power law regression model was constructed to correlate the SIF with input configuration parameters, such as internal pressure (P) and bending radius factor (BRF).

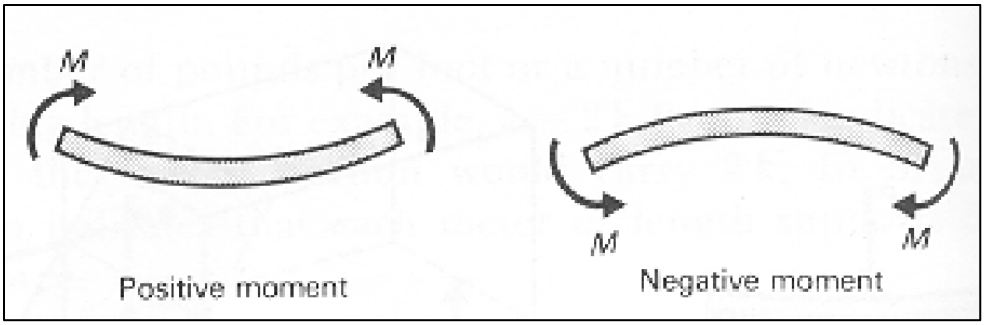


Figure 4-24. Bending moment sign convention-positive concave up, negative concave down

Through comprehensive analysis, it is found that the calculated SIF is independent of pipe size, as expected, and all the regression models can be generalized to a single power law function. The generic equation for the fitting is shown in Equation 4-14 and the regression coefficients are listed in **Table 4-4**.

$$SIF = aP^b BRF^c \qquad \textbf{Equation 4-14}$$

where SIF – Stress Intensification Factor

P – Pressure[psi]

BRF – Bend Radius Factor expressed as multiples of the pipe diameter

a, b, c – Regression Coefficients

Table 4-4. Regression coefficients for each configuration

Configuration	a	b	c
Pipe without Tee	4453.0	-1.0657	-0.744
Pipe with Tee Positive Bending	5310.7	-0.9654	-0.7251
Pipe with Tee Negative Bending	9364.1	-1.0757	-0.7789
Pipe with Tee Lateral Bending	3476.7	-1.0059	-0.5593

Figure 4-25 shows the relative severity of the bending configurations at an internal pressure of 45 psig. The maximum stress is the greatest for a lateral bend > positive vertical bending> negative vertical bending > pipe with no tee. To simplify the computation and application of the developed network, the regression coefficients for pipes with the tee under lateral bending are chosen to represent all three configurations, which can provide relatively conservative life prediction for Aldyl A pipes.

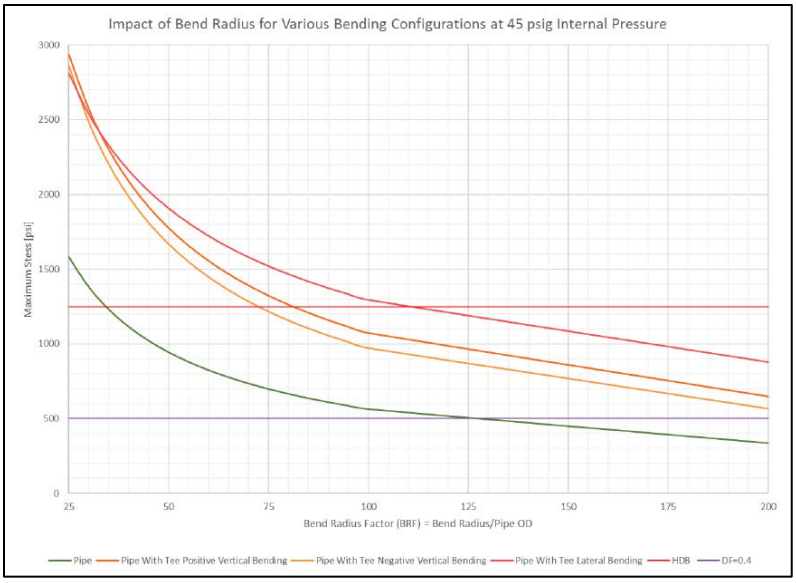


Figure 4-25. The relative severity of bending configurations at 45 psig internal pressure, SDR 11 pipe

The investigation of the soil movement effect on the pipeline has been performed in GTI project 21559 and 21584 Mobile Hybrid LiDAR & Infrared Sensing for Natural Gas Pipeline Monitoring. Gas distribution pipeline may experience high longitudinal strains in the events of soil movement resulting from external force, seismic activity, slope instability, flooding, and soil subsidence. The deformation of the pipeline may be either along the vertical direction because of soil settlement or along the horizontal direction because of landslides. The schematic representation of pipe deformation under vertical or horizontal soil movement is shown in **Figure 4-26**. It can be easily concluded that bending will be introduced on the pipe due to the combinatory effect of the vertical and horizontal soil movement. The rough estimation of the bending radius can be computed using the fundamental knowledge of solid geometry. The schematics of the resultant pipeline deformation are demonstrated in **Figure 4-27**.

(a) (b)

Figure 4-26. The schematic representation of pipeline bending undervertical soil movement (a) and horizontal soil movement (b)

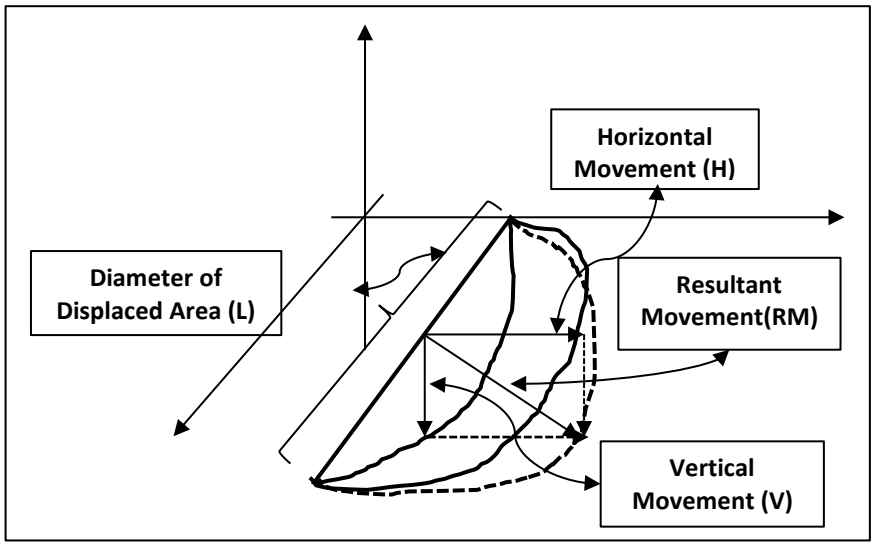


Figure 4-27. The resultant deformation of pipeline bending due to soil movement

As seen above, the bending radius (BR) can be calculated as

$$RM = \sqrt{V^2 + H^2}$$ **Equation 4-15**

$$BR^2 = (BR - RM)^2 + \frac{L^2}{4}$$ **Equation 4-16**

$$BR = \frac{RM}{2} + \frac{L^2}{4RM}$$ **Equation 4-17**

From the expression in Equation 4-14, it is trivial to find that the bending radius caused by soil movement can be easily incorporated in this formula. Considering the impingement, soil movement, and bending discussed above, the Bayesian network for operational SIF is shown in **Figure 4-28**.

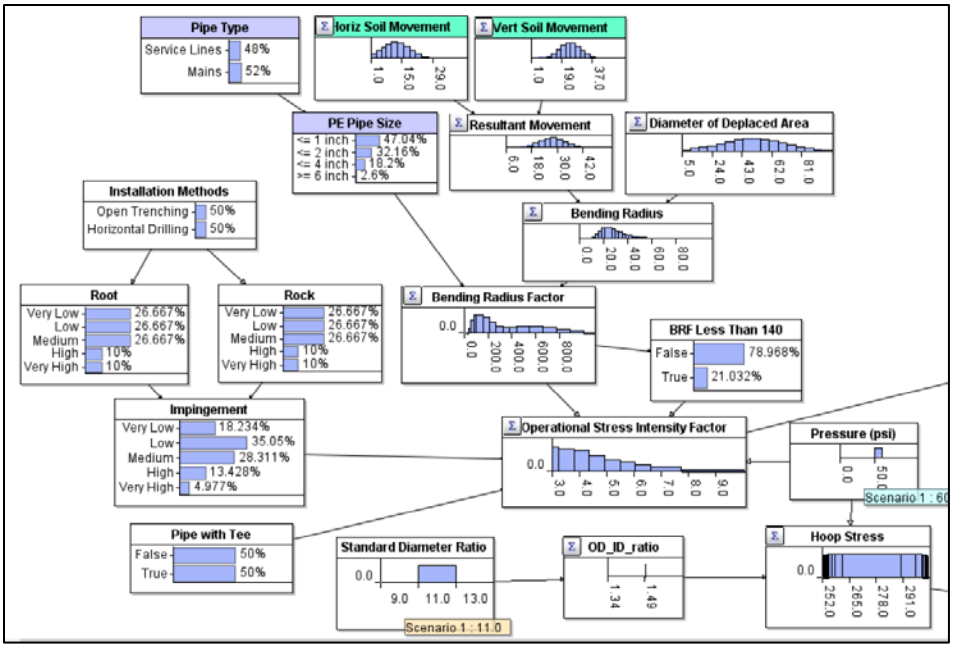


Figure 4-28. The Bayesian network for node "Operational SIF"

Risk Score

During the investigations carried out under the OTD 2.8.d (20649) project, it was noted that certain surface features visible in Cross-Polarized Light Microscopy (CPLM) and Scanning Electron Microscopy (SEM) appeared to correlate to times to failure in the RPM testing of pipe specimens. A categorical (multinomial) logistic regression model yielded surprisingly good predictive power. The model was later improved by incorporating relative ranking derived from Oxidation Induction Time (OIT) from thermal testing and Carbonyl Index (CI) from Fourier Transform Infrared (FTIR) testing. The developed logistic regression model was implemented in the Bayesian network to create the node probability table given the fitted model coefficients. Details for computing the probability of each state (i.e. outcome) are discussed below.

The multinomial logistic regression is a classification method that generalizes logistic regression to solve problems with more than two possible discrete outcomes[26]. Assume there are K possible outcomes, the multinomial logit model is to regress the $K-1$ outcomes against the pivot outcome, such as outcome K. The logit regression can be expressed as

$$ln(Pr(Y_i = 1)/Pr(Y_i = K)) = \beta_1 X_i \qquad \textbf{Equation 4-18}$$

$$ln(Pr(Y_i = 2)/Pr(Y_i = K)) = \beta_2 X_i \qquad \textbf{Equation 4-19}$$

$$ln(Pr(Y_i = K-1)/Pr(Y_i = K)) = \beta_{K-1} X_i \qquad \textbf{Equation 4-20}$$

where $\beta_1, \beta_2, \ldots, \beta_{K-1}$ are the model coefficient vectors corresponding to 1, 2, …. , $K-1$ outcomes, X_i is a vector representing the surface condition attributes of a pipe. Since the sum of the probabilities for K outcomes is one, the probability for each outcome can be computed as

$$Pr(Y_i = K) = 1 \Big/ (1 + \sum_{k=1}^{K-1} exp(\beta_k X_i)) \qquad \textbf{Equation 4-21}$$

$$Pr(Y_i = 1) = exp(\beta_1 X_i) \Big/ (1 + \sum_{k=1}^{K-1} exp(\beta_k X_i)) \qquad \textbf{Equation 4-22}$$

$$Pr(Y_i = 2) = exp(\beta_2 X_i) \Big/ (1 + \sum_{k=1}^{K-1} exp(\beta_k X_i)) \qquad \textbf{Equation 4-23}$$

$$Pr(Y_i = K-1) = exp(\beta_{K-1} X_i) \Big/ (1 + \sum_{k=1}^{K-1} exp(\beta_k X_i)) \qquad \textbf{Equation 4-24}$$

Given the derivation above, the node probability table of the "Risk Score" node can be constructed using the AgenaRisk API. The completed subnetwork is illustrated in **Figure 4-29**.

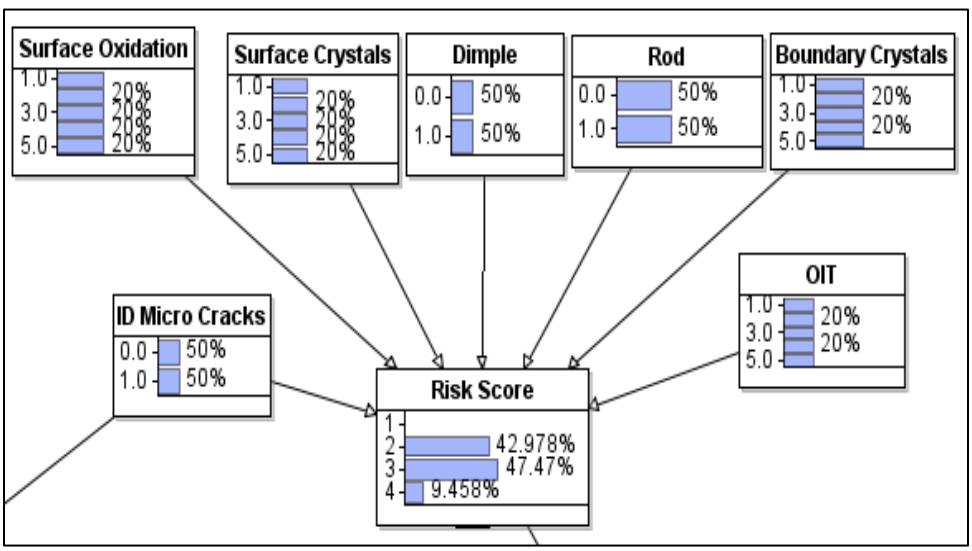


Figure 4-29. The Bayesian network for node "Risk Score"

Sub-networks Assembly

In this part, the sub-networks developed above are integrated as an overall Bayesian network as shown **Figure 4-30**. Using the developed network, on the one hand, the remaining useful life (RUL) of pipelines can be predicted given context conditions, such as buried

environment, application temperature, and pipe inner diameter surface features. On the other hand, the context conditions can be diagnosed based on the evidence for pipeline RUL. The forward and backward reasoning process will be provided later to validate the integrity of the developed risk assessment network.

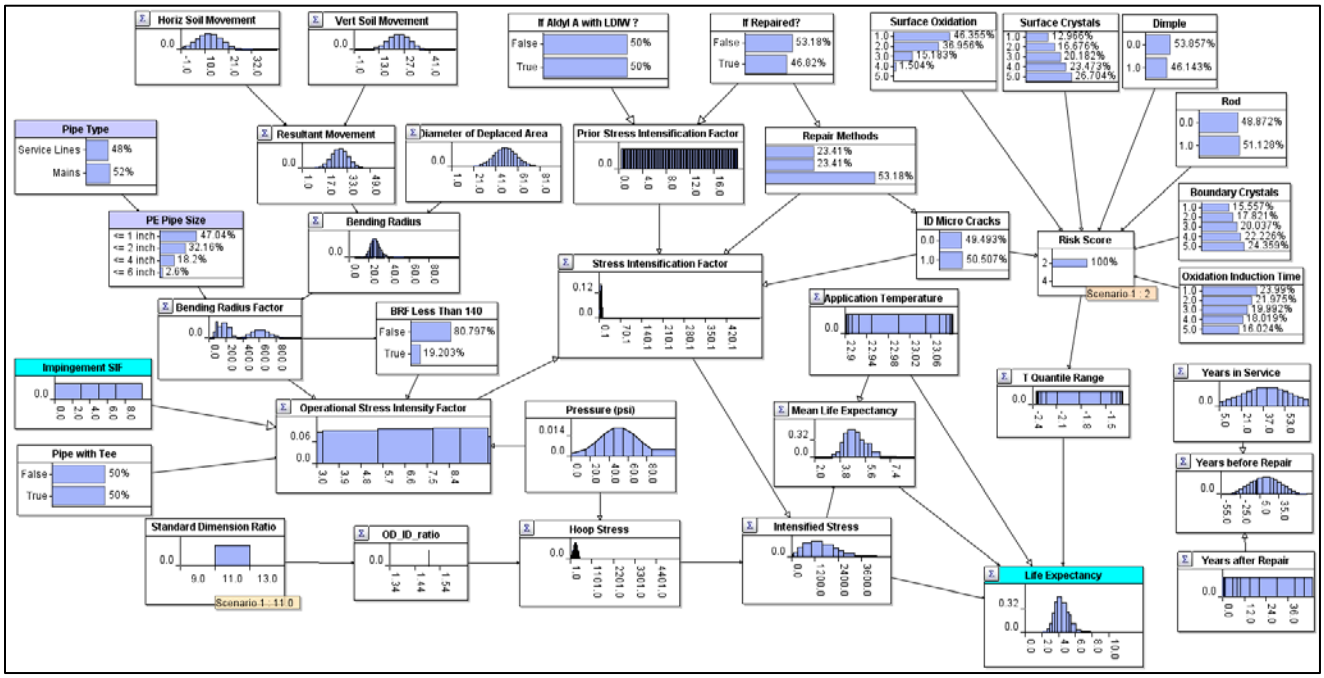


Figure 4-30. The integrated Bayesian network for pipeline risk assessment

Bayesian Network Testing and Validation

Risk Score Validation

For Bayesian network, one node is independent of the others if the evidences for all of its parent nodes are provided. Analogously, in the developed overall network, the risk score can be computed given its parent nodes representing the inside diameter surface features. In the OTD project (4.12-GTI 21301), different combinations of surface features and the predicted risk score from logistic regression were listed. Plugging these data into the developed network, the risk score for each scenario should be replicated. The risk scores from the original regression model and Bayesian network are compared in **Table 4-5**.

Table 4-5. The risk score comparison between logistic regression and Bayesian network

Risk Score from Regression	Probability of Each Category				Risk Score from network
	1	2	3	4	
3	7.40E-05	0	0.998062	0.001864	**3**
4	0.000179	0	0.010156	0.989665	**4**
3	0.000707	5.00E-06	0.998109	0.001179	**3**
3	0.001328	0.496864	0.501255	0.000553	**3**

3	0.001982	0.03762	0.958174	0.002224	**3**
2	6.00E-06	0.972903	0.027091	0	**2**
3	0.001328	0.496864	0.501255	0.000553	**3**
2	0.000366	0.982811	0.015212	0.00161	**2**
2	0.00032	0.993946	0.003589	0.002146	**2**
2	4.00E-06	0.997916	0.00208	0	**2**
3	0.002037	0.494501	0.501623	0.00184	**3**
3	0.001875	0.012431	0.984085	0.001609	**3**
3	0.000811	0.00141	0.997339	0.00044	**3**
3	7.60E-05	0.001317	0.998605	2.00E-06	**3**
3	0.002037	0.494501	0.501623	0.00184	**3**
2	2.00E-06	0.999757	0.000241	0	**2**

From **Table 4-5**, it can be seen that the developed network can completely reproduce the results from the original logistic regression model when the category is determined by the maximum predicted probability. In the developed Bayesian network, the probability for each category will be considered for risk assessment, which will minimize the information loss from the logistic regression model.

Rate Process Model Validation

In this Bayesian network, the pipes are undergoing slow crack growth failure pattern under the normal service condition. The risk score computed above was initially generated relative to the reference SCG model. Different scores mean different levels of risk for SCG failure. In order to validate the reference SCG model, the Bayesian network is run multiple times with different stress and risk scores. The prediction limits (PL) from the SCG reference model and the Bayesian network are plotted in **Figure 4-31**.

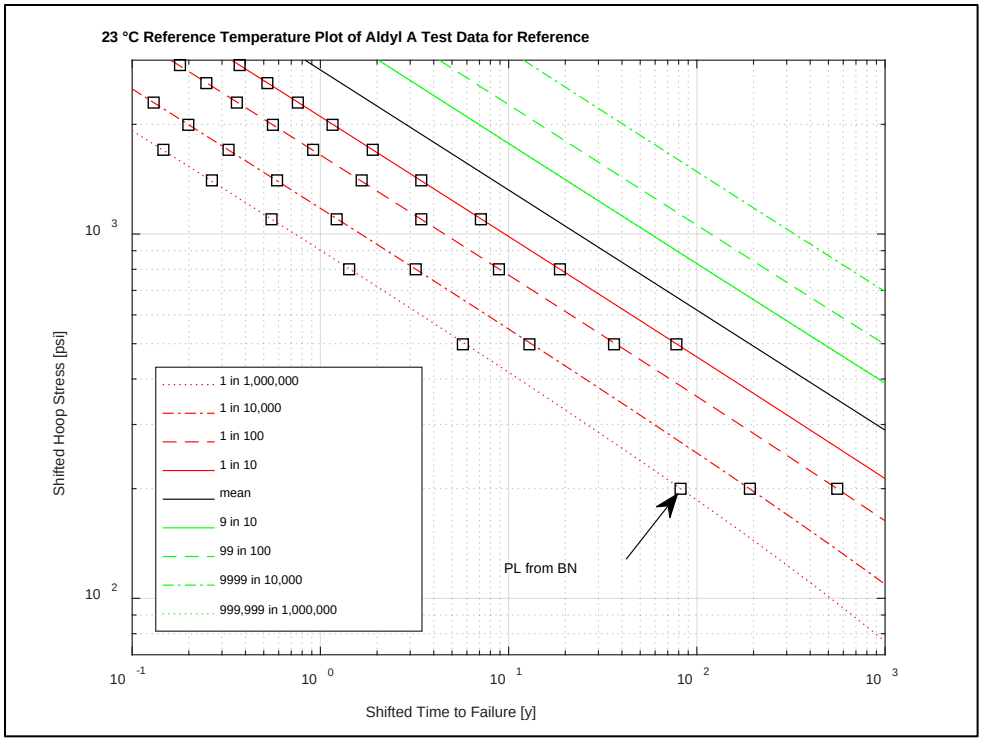


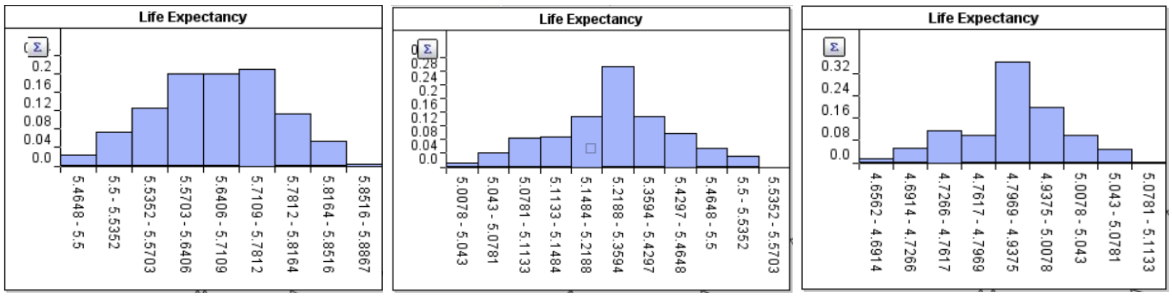
Figure 4-31. The reproduced prediction limits using Bayesian network

Remaining Useful Life Prediction

The critical application for this network is to assess the risk of the pipe by predicting its RUL given specific context conditions, such as temperature, buried environment, inner diameter surface features, and so on. If the pipe is under the same internal pressure, the RUL will be decreasing with higher risk score. To validate this conceptual understanding, three scenarios are generated and listed in
Table 4-6.

Table 4-6. The context conditions for 3 scenarios

Scenarios	Intensified Stress (psi)	Temperature (Celsius)	Dimple	ID Micro Cracks	Rod	Boundary Crystal	Surface Crystal	Oxidation Induction Time	Surface Oxidation	Risk Score
3	500	23	1	1	1	1	1	4	4	1
1	500	23	1	1	0	5	2	1	2	2
2	500	23	1	1	0	1	2	3	5	3



(a) Risk Score=1 (b) Risk Score=2 (c) Risk Score=3

Figure 4-32. The final distributions for the log of the life expectancy (hrs) Mean life=53.1 years, (b) Mean life=21.7 years, (c) Mean life=8.7 years

To investigate the effect of prior SIF, 10 scenarios are generated by changing the evidence for nodes "If Aldyl A with LDIW", "If Repaired", "Repair Methods", "Years in Service" and "Years after Repair". Using the network, their corresponding median and quantile life prediction can be easily computed. It should be noted that the life prediction may not represent the true life for the pipe, because the overall Bayesian network has not been calibrated using true testing data for pipes undergoing the combination of different loading conditions. However, the trend for the life prediction should be consistent with the conceptual understanding of the pipeline failure mechanism. Those assumed 10 scenarios are summarized in **Table 4-7**, in which the first row listed the node names. The life prediction results are summarized in **Table 4-8.**

Table 4-7. The context conditions for 10 scenarios

Scenarios	If Aldyl A with LDIW	If Repaired	Repair Methods	Years after Repair	Years	Temperature	Pressure	Bending Radius Factor
1	False	False	NA	0	40	23	60	150
2	False	True	Butt Fusion	10	40	23	60	150
3	False	True	Butt Fusion	20	40	23	60	150
4	False	True	Butt Fusion	30	40	23	60	150
5	False	True	Mechanical Coupling	30	40	23	60	150
6	True	False	NA	0	40	23	60	150
7	True	True	Butt Fusion	5	40	23	60	150
8	True	True	Butt Fusion	10	40	23	60	150
9	True	True	Butt Fusion	15	40	23	60	150
10	True	True	Mechanical Coupling	15	40	23	60	150

Table 4-8. The median and quantile life prediction for those 10 scenarios (years)

Scenarios	Median	Upper Quantile	Lower Quantile
1	35.54	89.92	15.22
2	22.17	41.27	8.21
3	21.50	39.40	7.99
4	17.16	39.15	7.70
5	13.34	28.95	6.14
6	23.41	43.76	9.11
7	10.48	24.05	4.77
8	10.15	22.01	4.65
9	10.69	24.25	4.96
10	8.52	17.86	4.01

From above table, it can be seen that the predicted life will be shorter for the pipe that has been repaired longer time back. The reason is because it is assumed that the plastic pipe will be squeezed once it is repaired. The mechanical coupling may also decrease the life time and the reason is the localized bending will be introduced around the coupling edge. Those preliminary assumptions can be updated once additional knowledge is acquired.

Conclusions and discussions

This section presents the procedures of designing, developing, and validating the probabilistic risk assessment model. Bayesian network is proven to be a good candidate for expressing the semantic ontology of pipeline risk phenomenon. The inherent causal relationship could greatly reduce the complexity and scale of the network model. By processing the Aldyl A hydrostatic testing data sets, the effects of years in service, squeeze-off and LDIW are successfully taken into account as the prior SIF distribution. Impingement induced SIF is modelled by combining the theory of plates and stress strain curve considering large plastic deformation, the sub-network developed from which is also calibrated and corrected using the hydrostatic testing results. Each sub-network can be continuously refined with additional data acquisition. Then, they are assembled to compute the life expectancy of pipeline segments given different context conditions. The risk model can be combined with smart data collection, pipeline inspection, and mitigation practice to provide informative guidance for integrity management and concurrent situation awareness.

It should be noted that the contribution of each sub-network to the final failure time is determined by the investigator's best knowledge. Additional research is needed for the interactive nature of different sources of SIF. The accuracy of each part of the network may also be further updated if more reliable data is collected.

Section 4 Bibliography

19. Pearl, J., *Causality: Models, Reasoning, and Inference* 2000: Cambridge University Press.
20. Fenton, N. and M. Neil, *Risk assessment and decision analysis with Bayesian networks*. 2012: CRC Press.
21. Haine, S., *Hazard Analysis & Mitigation Report on Aldyl A Polyethylene Gas Pipelines in California*. 2014, California Public Utilities Commission.
22. Bianca Pinheiro, I.P.a.S.B.C. *Stress Concentration Factors of Dented Pipelines*. in *6th International Pipeline Conference*. 2006. Calgary.
23. Keating, A.J.R.a.P.B., *Stress Concentration Solution for a 2D Dent in an Internally Pressurized Cylinder.* JOURNAL OF ENGINEERING MECHANICS, 2007. **133**(7): p. 792-800.
24. Godoy, L.A., *Thin-walled structures with structural imperfections*. 1996, New York: Pergamon.
25. Jajo, J.O., *Dent behaviour of steel pipes under pressure load*. 2014, Electronic Theses and Dissertations.
26. *Multinomial logistic regression*, in *Wikipedia*.

5. Structured Light Method for Internal Inspection of Aldyl-A Pipe

Introduction

Identification and classification of the current vintage pipeline inner wall damage precursor are of critical importance. However, there is a limited success in sensing and characterizing the long lengths of deteriorating plastic pipelines with a high probability of detection (POD), and the capability of the currently available accelerometers and imaging technologies that can be miniaturized and integrated into smaller CTS size pipes for a fast scan is questionable and needs a systematic assessment. There are several known inner pipe wall surface conditions that increase the probability of premature failure due to slow crack growth. The major drivers of premature failure due to slow crack growth are bending stresses due to tight bend radii, impingement, and fittings. Damage is also introduced by pipe squeeze-off during maintenance operations. These conditions identification has been investigated by various nondestructive evaluation (NDE) techniques, such as direct visual/optical methods using CCD cameras, ultrasonic testing, liquid-coupled acoustic measurement (e.g. sonar) and laser based surface inspection approaches including light detection and ranging (LiDAR) and laser topography[27-29]. However, these currently available technologies suffer either from low-sensitivity and resolution for small damage precursors, or complex settings and large system footprint that makes it incapable for plastic gas distribution pipes with much smaller diameters. In this task, led by the PI at MSU, we developed a multi-spectral and miniaturized optical sensing platform for imaging the smaller CTS size pipe inner walls with high sensitivity and specificity of damage identification. The literature review and progress in the 1st quarter are presented in the following sections.

Principle of 3D optical inspection system

3D profiling has been widely used and has a lot of applications in manufacturing and medical field. The principle of capturing 3D image is mathematically based on the triangulation. This technique depends on two basic and critical components in the system, which are camera and projector. The optical projector/source sends structured light to the object, where a typical structure light pattern is shown in **Figure 5-1**. The benefit of using this light is to computationally capture the depth information from the 3D object from a relatively simple light projection. In the literature, a pinhole model is used to capture the light that reflected from the object.

$$\widetilde{m}^c \propto K^c[R^c\ t^c]\widetilde{M}^w \qquad \textbf{Equation 5-1}$$

where $\propto$ stands for the equality up to scale factor, $[R^c\ t^c]$ is extrinsic parameters matrix and K^c is the camera parameters whose matrix form is shown in eqn. **Equation 5-2**:

$$K^c = \begin{bmatrix} \propto & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$ **Equation 5-2**

u_0 *and* v_0 are coordinates of the principle point, while $\propto$ *and* β are the focal lengths.

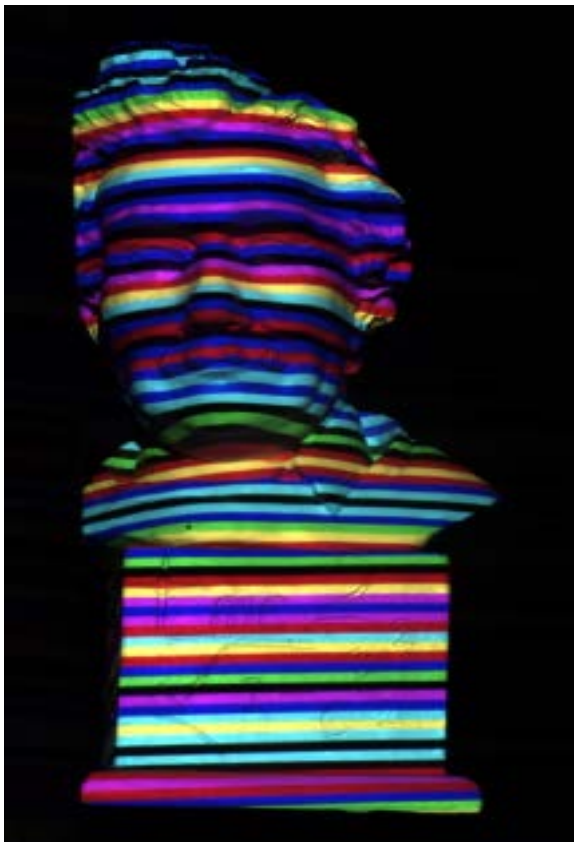


Figure 5-1. Example of a structured light pattern projected onto an object [30]

By using the structured light illumination, the collected images by the camera will be reconstructed into 3D data. Figure 5-2 illustrates the triangulation system by a camera and light source. In this task, by taking advantage of the endoscope technique, the innovative imaging principle of using the multicolor multi-ring is considered to increase the resolution and sensing speed for possible in-line inspection (ILI) integration. Although there are many kinds of structured light methods, Single Shot (SS) technique is one of the most used techniques, and we will adopt it for the first prototype development. Since the single shot just needs one image of the scene to reconstruct 3D data, it will simplify the sensor hardware design as well as reduce the image processing load. In other literature in the medical field, one design of structured light endoscope needs a camera and a slide projector to collect the 3D data, which will also be investigated in this task. Figure 5-3 shows the endoscope technique using for clinical applications at the diameter of 3.8mm.

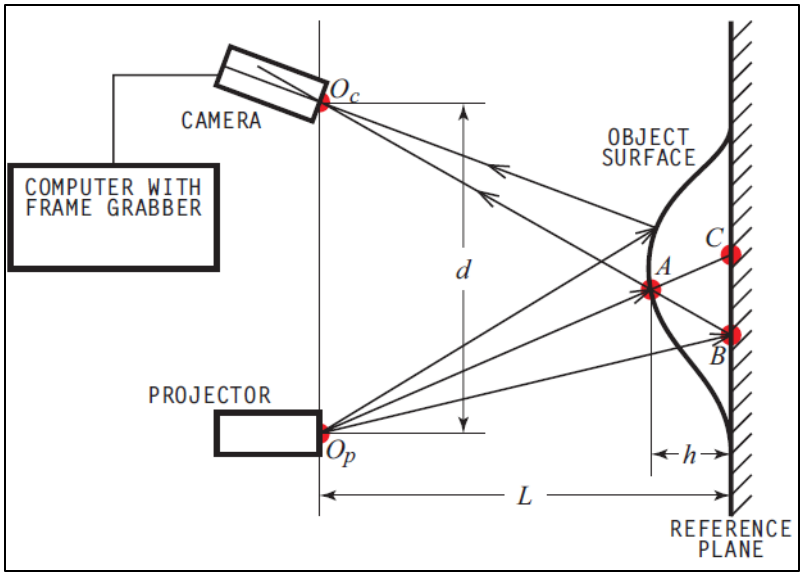


Figure 5-2: Illustration of a 3D structured light measurement [31]

Optical sensing is one of the NDE techniques that have been improved significantly in the past few years. There are several ways under optical sensing to produce a 3D image. However, two different techniques will be the main categories which are passive acquisition and active acquisition, in which laser and structured light will be classified under active acquisition. The following chart shown in **Figure 5-3** illustrates the main categories of the optical sensors.

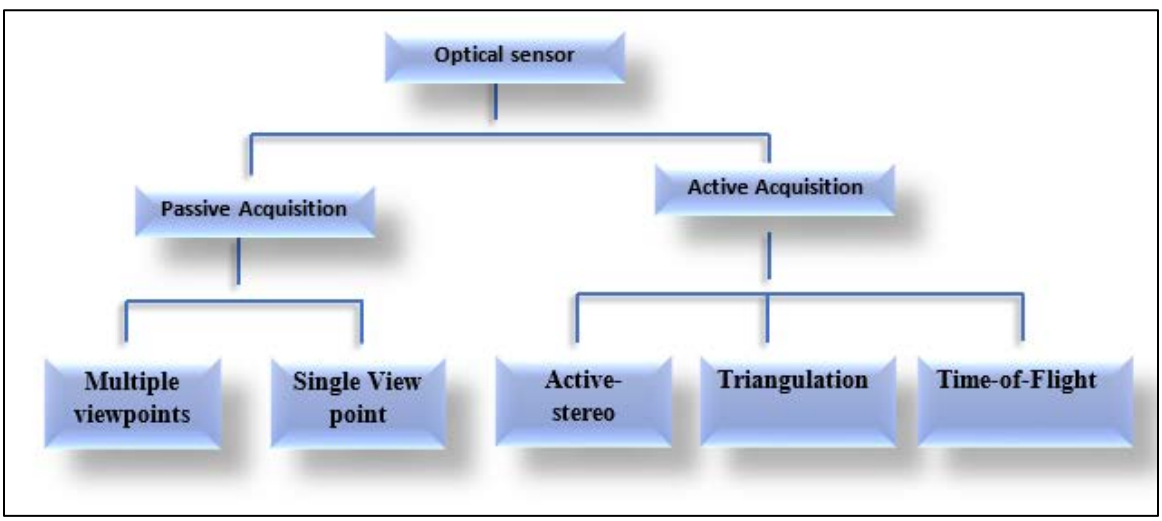


Figure 5-3. Classification of the optical sensors category

Passive Acquisition

Passive Acquisition is one of the methods that can be used to obtain 3D optical images. This technique is called passive because the data is collected by using the camera without using any active source, such as lasers to provide strip line or by using structured lights. Multiple

viewpoints and single view point techniques are classified to be the main methods for the passive acquisition

Multiple viewpoints

This technique is used to take the image from different viewpoints/angles. Two cameras or more can be used to collect the data from different angles. By using multiple cameras, the technique will be called stereo vision. There are different kinds of stereo with a different number of the cameras that have been used. For example: if two cameras have been used to capture the image the technique, is called binocular stereo. Similarly, trinocular stereo is when three cameras have been used to capture the image. Also, if a single camera has been used to take the image for the same point but from different locations and time, this technique is called structure from motion. By processing the images that have been taken from the different locations or different cameras, the 3D rays will be determined. Finally, from the 3D rays, the 3D position of the point in the scene can be determined. One illustration of the multiple viewpoints is shown in **Figure 5-4.**

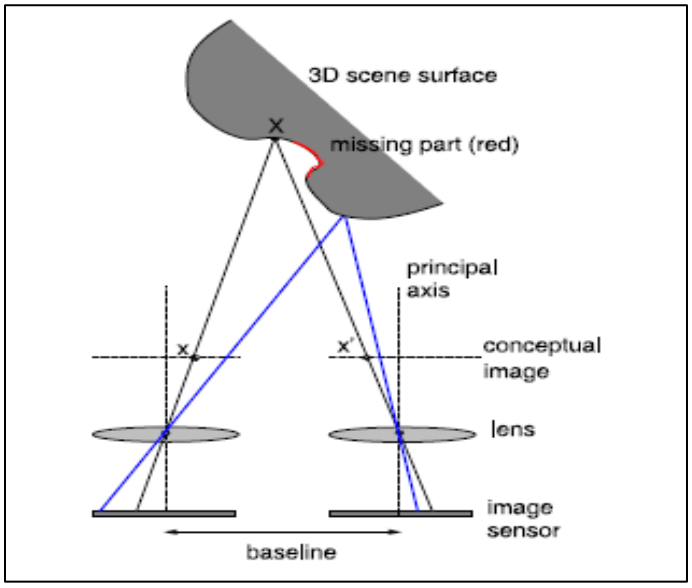


Figure 5-4. Illustration of multiple viewpoints technique [31]

Single view points

In this technique, the captured image does not come from multiple cameras or the camera motion. The captured image, in this case, will be taken by using the object details, for example, the texture of the object, shading, or focus, etc. **Figure 5-5** shows that the 3D image has been generated by shape from shading. This technique depends on the reflection from the object. The pixels in the reconstructed image illustrate the intensity of the reflection that comes from the object shade, and by using regularized surface fitting, the 3D image can be reconstructed. Overall, the multiple viewpoints method is more accurate and more efficient

than single view point method. This technique's limitations and the enhancement of the reconstructed image are described in [32]. Shape from shading technique is known as Photometric stereo. Main steps to generate a 3D image by using this technique is that taking more than one image for the same point but the illuminations for the scenes are different. The other technique is called shape from focus. Taking two images from different depths of field is the main idea of processing 3D image, and the approach is described in [33].
In summary, single view technique is not as good as multiple viewpoints technique in terms of speed and regulation. Therefore, multiple viewpoints technique is more commonly used. However, the passive acquisition is not suitable for this project. The active acquisition will be the tool for making the 3D imaging which will be reviewed and discussed in the following section.

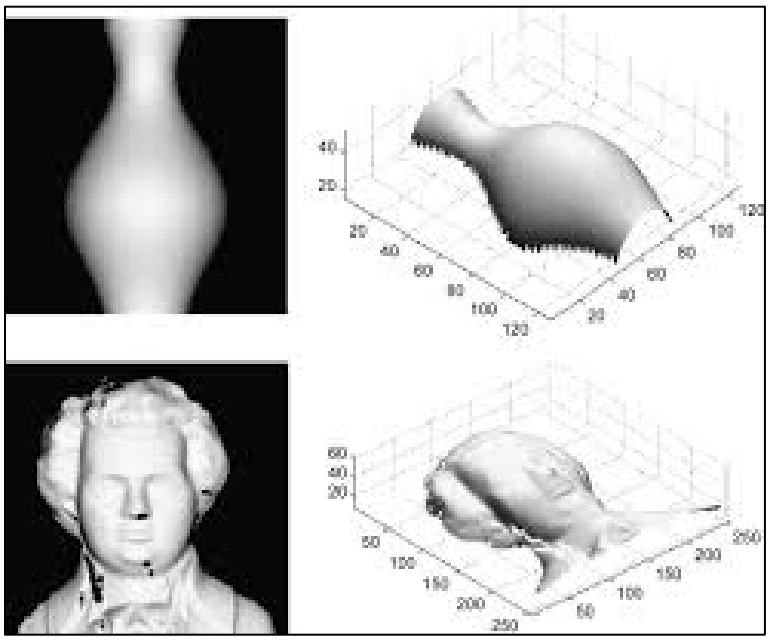


Figure 5-5. Illustration of generating 3D image by using shape from shading [28]

Active Acquisition

Active Acquisition has different techniques to capture the viewpoints comparing with passive techniques. As illustrated in the passive techniques, captured image does not need structured light or laser stripe to collect the data. However, in Active Acquisition, optical detectors such as camera need to be used to detect the point. In addition, the active sources such as laser source or structure light should be considered to complete the imaging process.

Active-stereo

Active stereo has a similar concept with multiple viewpoints. In this technique, a light source with special features is used to replace the function one of the cameras in the multiple viewpoints method. **Figure 5-6** illustrates the principle of the Active-stereo method with laser scanning. The light source has been focused on the scene; then the camera captures the

view. There are different kinds of the light sources. Structured light and laser are two of the most known sources. These techniques will be discussed in the following sections. Active stereo is more common than passive techniques because of the high resolution, and the data that have been taken by the laser or structured light are more accurate. However, to generate a 3D image by using a camera and light source, the triangulation should be considered to calculate the depth of the scene, which makes the method mathematically rigorous and challenging.

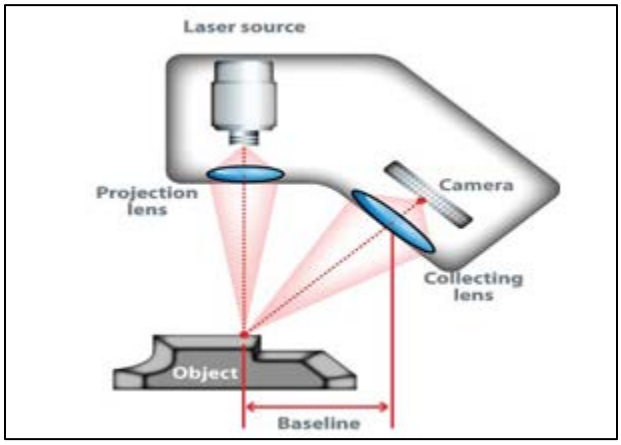


Figure 5-6. Illustration of active stereo with laser scanning technique [31]

Laser Scanning

A lot of applications have used laser scanning to get a 3D image. Laser scanning is using a strip of laser light, or it can be a circle-shaped source to scan the inner surface of an object, (e.g. pipelines in this project, which will be discussed in detail later). Basically, a camera with high resolution has been used as a sensor to capture the scene. Both the camera and the laser source work as one unit to produce a 3D image. **Figure 5-7** illustrates how the laser and camera work together. The distance between the camera and the laser source and the illumination is calculated by using triangulation to create the 3D images. When the laser projects the light on the object, the laser light will capture the object shape. Then by knowing the camera and laser positions, the depth of the object can be easily calculated.

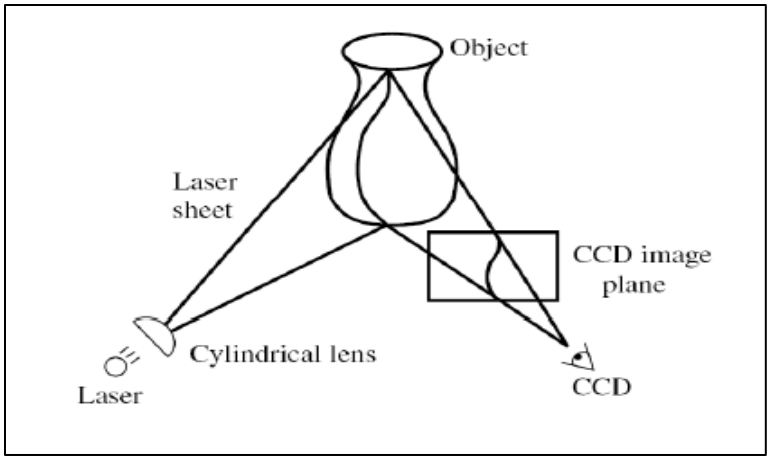


Figure 5-7. Laser scanning technique using CCD camera and laser source [34]

Structured Light

As discussed in the laser scanning section, structured light is similar to the laser technique in setup, however, a structured light projector is used in this technique instead of the laser source. The structured light technique has been used for a lot of applications in industrial and medical testing. The principle of structured light is shown in Figure 5-8, which depends on the relative positions of the camera and projector. In this system, the projector is connected to the computer and is controlled to generate pre-designed patterns. There are different kinds of patterns that are used for structured light such as: Spatial patterns, Temporal patterns, Colored patterns.

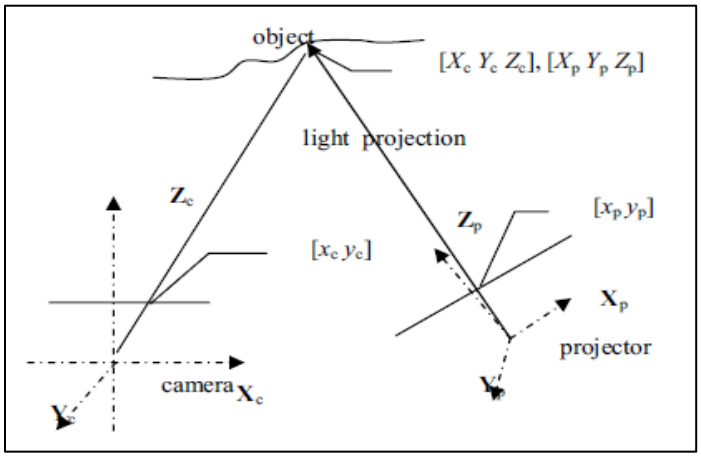


Figure 5-8. structured light technique with a light projector and camera [35]

Multi colors multi rings approach:

Laser ring approach is a robust technique to scan surfaces, but the main drawback is its long scanning time due to the small number of 3D points acquired from each single frame. Another factor is the limitation of the spatial resolution due to the width of projected laser

ring. In this section, we present a multi rings approach to increase the number of acquired 3D point from each recorded frame and to increase the resolution of the acquisition system. In this scenario, a 1D stripes sequence or a 2D grid is projected to the object to be scanned. The projection of multiple scanning lines leads to the rise of the correspondence problem (i.e. the correspondence between the projected and imaged scanning lines). Different methods are used to solve this problem. It varies from shifting of binary patterns to the use of frequency domain methods. Experiments show that the methods that are implemented with multiple projections provide dense 3D acquisitions but with the cost of processing time while single shot techniques provide nearly real-time results but provide a more sparse 3D grid that depends on the number of projected stripes in each single frame[36, 37]. In the case of the pipe internal surface scanner, the camera is moving inside the pipe which eliminates the possibility of using multi images technique due to the difficulty of acquiring these images for the same scene during the camera movement along the pipe. Thus, a single shot system is employed to account for the effect of scanner movement.

Colored single shot patterns:

In this method, multiple lines with different colors are projected to the surface with special sequences in order to acquire them uniquely[38, 39]. This method requires only one frame to perceive the depth of the scanned object. The method produces surfaces with less resolution due to the limitations on the number of projected lines, but it is considered fast and can be adapted to suit real-time systems and the imaging of moving objects [30].

Projector pattern

Single shot techniques require the projection of special patterns where each set of sub-patterns are unique within the projected color sequence to retrieve the actual location of the projected colored lines. For this purpose, the sequence is required to have a set of words that are uncorrelated with their entire projected sequence. De Bruijn sequences offer such capability by providing a set of uncorrelated sets of numbers. Each set is governed by two factors k and n. Where k represents the number of values that can be chosen for sequence elements, and n represents the length of each sub-pattern. These factors also govern the total length of the sequence where the total length equals k^n.

Reconstruction algorithm

The reconstruction algorithm is executed in the following sequence:

Image enhancement

The recorded images are passed through a number of steps that may include noise reduction, color correction and transformation to another color spaces, etc. these steps enhance the image quality and improve the image segmentation of the projected rings.

Segmentation:

One of the main operations in the line detection is the segmentation. Each layer is segmented by taking the gradient of the image with Sobel filter. A 2D filter is used because the lines are circular and the color change happens in both the vertical and horizontal axes. After finding the gradient in each channel, they are combined by using an ***or*** operator:

$$final\ edge = edge_r | edge_g | edge_b$$

The gradient is thresholded to separate the foreground from the background.

Color sequence detection:

In this step, the algorithm compares the sequence of the projected and detected pattern by comparing the values along the projection axis. In the case of rectangular pattern projection, the comparison is made by exciting a raster scan along the horizontal or vertical axis. In the case of the pipe scanner, the color change detection is done along the radial axis as explained by the direction of the arrows in **Figure 5-9**.

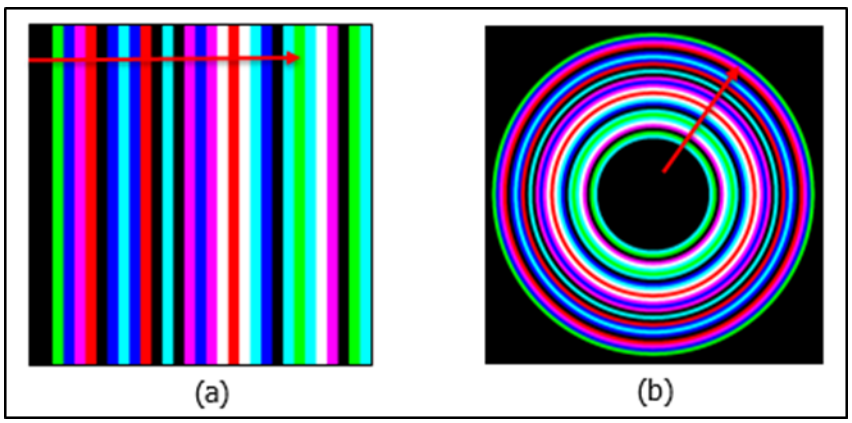


Figure 5-9. Projection patterns: a) rectangular, b) circular

The algorithm is utilized to check the color difference instead of checking the absolute value of the color because it is more robust to noise due the luminance differences [30]. The algorithm checks the detected sequence for any potential matching with projected sub patterns. For example, if the word length is four, then the algorithm starts matching the first four detected edges with the first four projected colors. If there is no match, the algorithm shifts the detected sequence and repeats the comparison process. If there is no match the projected colors windows is shifted and the comparison process is repeated. If a match is found, the four detected edges are registered as true edges, and the depth is extracted. Figure 5-10 shows the comparison process between the two hypothetical channels where the first one is the projected channel while the second is the received channel. The algorithm selects four elements from the projected sequence and starts shifting through the received channel

until it finds a match as shown with the sequence [-1 0 0 1]. In our case, the matching is executed through the 3 RGB channels at the same time.

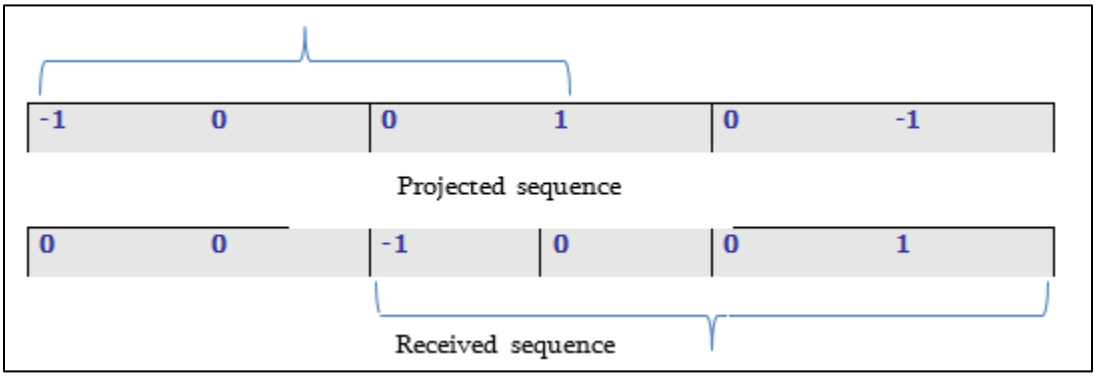


Figure 5-10: Matching process

Experimental

Defects and cracks in the pipes

In the beginning, 3 inch diameter pipes are considered as the specimen. Figure 5-11 shows the pipes that have been used in the experimental work.

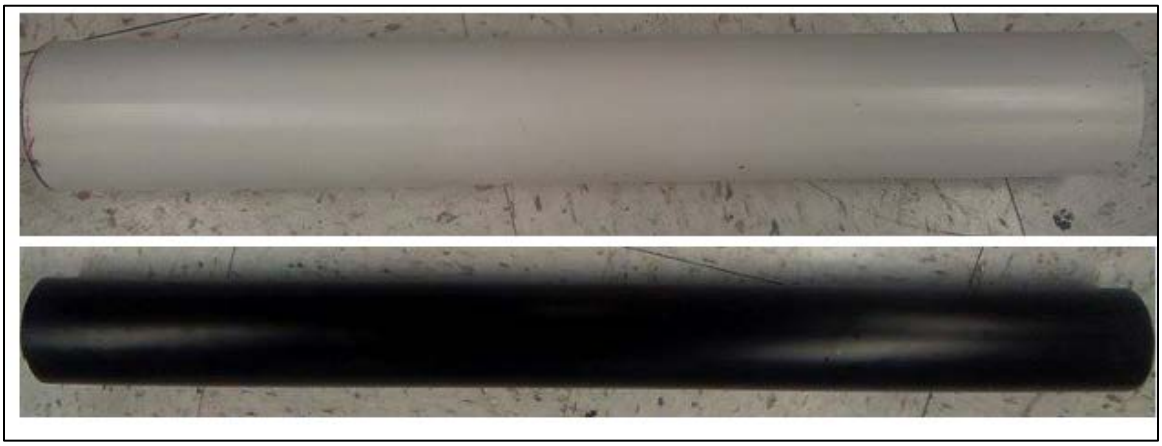


Figure 5-11. Pipes for the experimental work

Then some defects were introduced into the pipe as illustrated in Figure 5-12. Through hole defects with different diameters were introduced. In the white pipe specimen, the smallest hole in the pipe has the diameter of 0.0787 inches, and the biggest hole is 0.314 inches. Also, there are additional holes with 0.157 and 0.197 inch diameters. Moreover, some screws were installed to mimic different defect types. In the black pipe specimen, some damage was introduced to simulate different kinds of defects. **Figure 5-13** shows cracks with different directions on the pipe wall with a crack width 0.039 inches.

Figure 5-12: The holes that have been introduced as defects

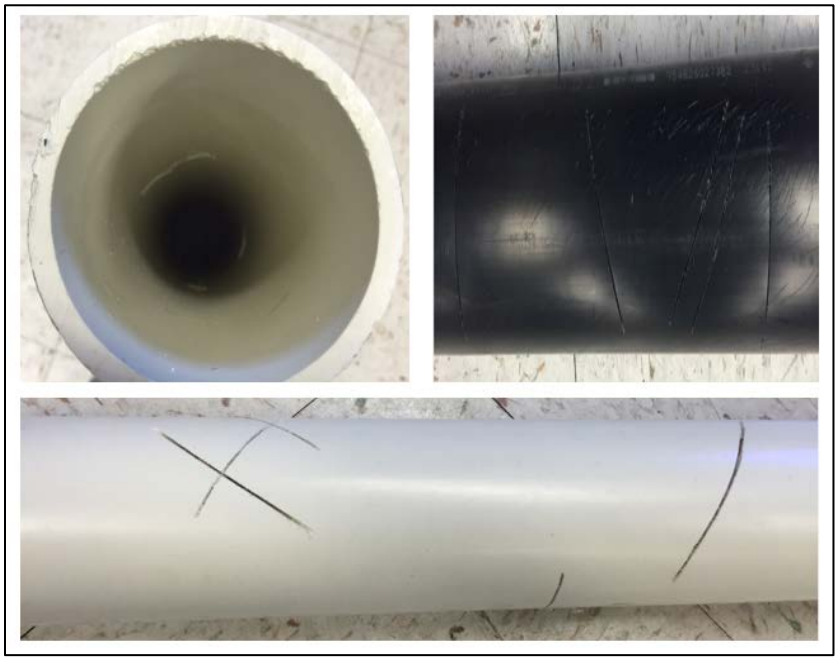


Figure 5-13. The linear cracks in the spacemen

Because our focus is to minimize the sensing system to fit the specimen smaller than 1 inch, a smaller pipe provided by GTI with different kind of defects was shipped to LEAP and tested in Q2. The diameter of this pipe is 1.85 inches. **Figure 5-14** shows the second type of pipe that has been used in the experimental work in Q2.

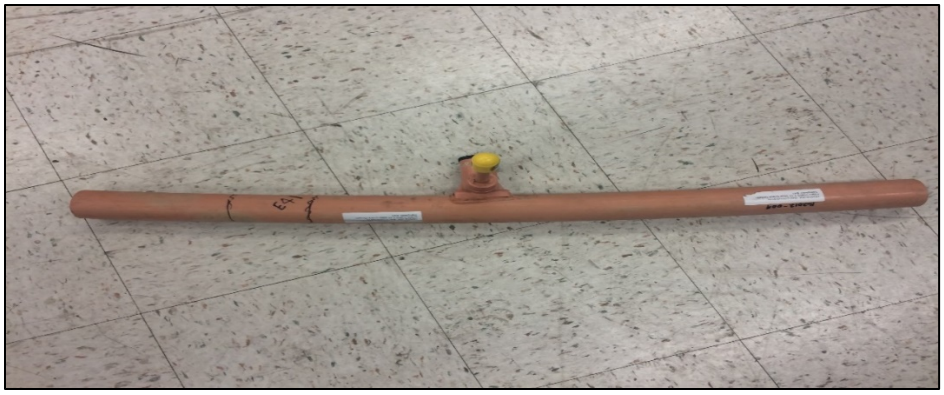
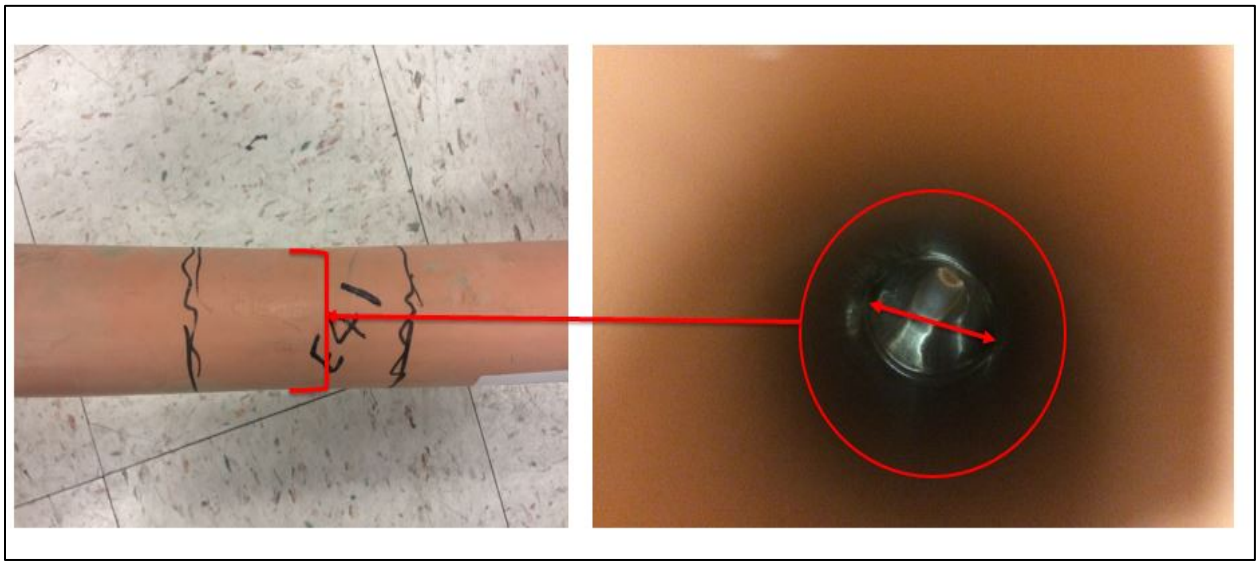


Figure 5-14. Pipe with a small diameter (1.85 inches) and damage section

The damaged section of this pipe is explained in **Figure 5-15** with **Figure 5-15.a** showing the outer section of the damaged area while **Figure 5-15.b** is showing an internal view that was taken by an endoscopic camera.



(a) (b)

Figure 5-15. Optical outer and inner view of the E41 damage of GTI pipe sample damaged area

Scanning without light source

Before performing the laser scanning or structured light scanning, passive scanning is considered to show the defects without any light source. Two different cameras were used to scan the inner side of the pipes as shown in **Figure 5-16**. These cameras have different diameters. The diameter of the smallest camera is 0.23 inches, and the diameter of the largest camera is 0.62 inches. The camera is installed in a special prototype that was produced by 3D printing. After installing the camera and moving it inside the pipe, a video is recorded to capture all inner surfaces in the pipe. **Figure 5-17 a** illustrates the cracks that have been

captured by using the two camera. Also, **Figure 5-17 b** shows the holes that were introduced into the pipe.

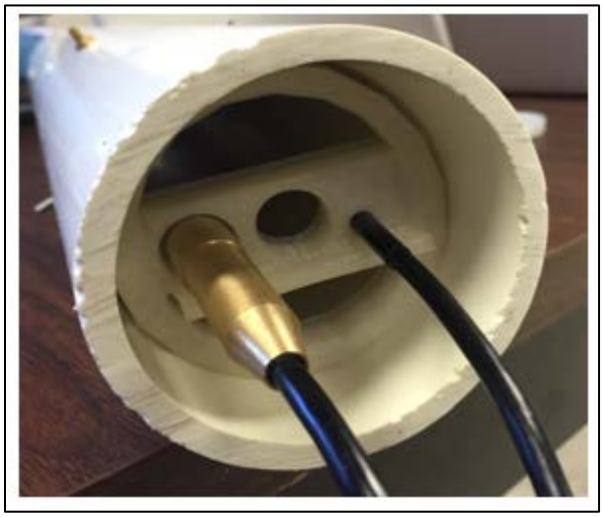
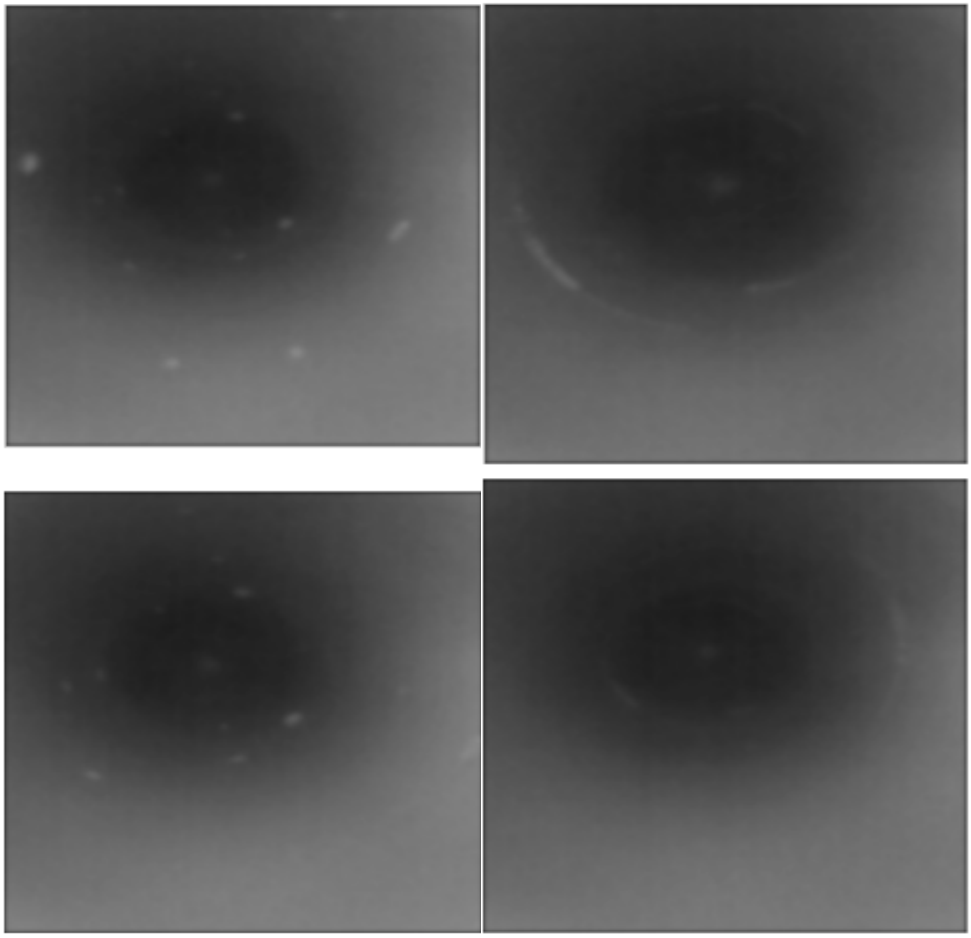


Figure 5-16. Two different cameras installed inside the pipe



(a) (b)

Figure 5-17. The different defect types captured by using the camera

Using a single passive camera gives an initial indication about the existence of damage on the pipe's surface but it doesn't provide enough information about the size and the shape of the damaged area. This is due to the lack of any depth information in the recorded camera frames.

Single ring scanning

In this part of the research, a single ring laser is considered to be the light source of the scanner. Laser source with a specific red ring was used to detect the cracks and the deformation. An image of the laser source and the projected red laser ring is shown in **Figure 5-18**. The relation between the diameter of the ring and the distance between the ring and the laser source is (1:1). This means that if the distance between the laser and the scene is 1 meter, then the diameter of the ring on the scene will be 1 meter.

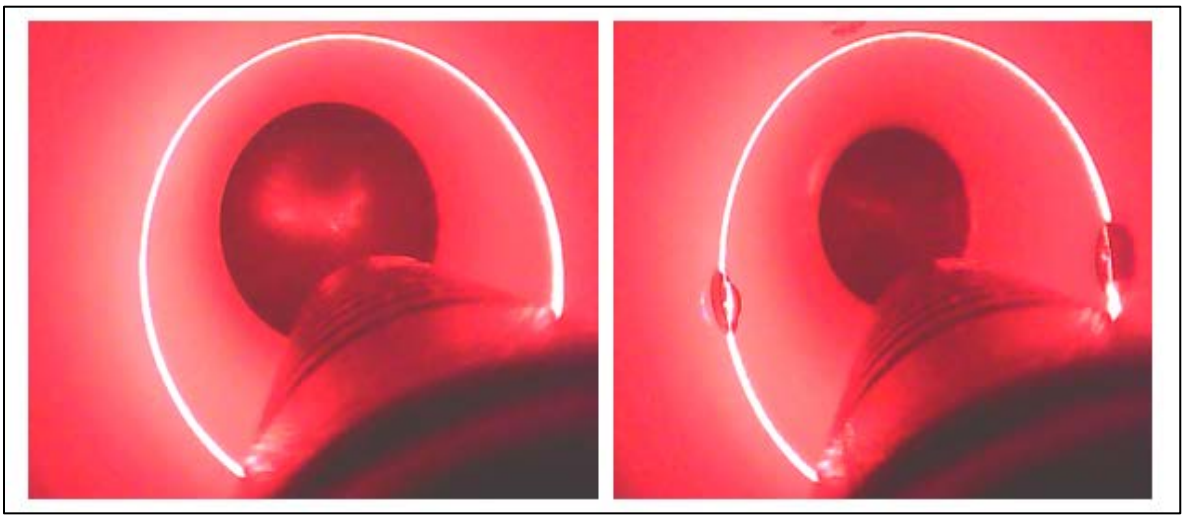
Figure 5-18. The laser source used in the sensor prototype

Scanning with small view angle camera

One of the cameras that was shown in **Figure 5-18** was used with the laser source to scan the pipe. **Figure 5-19** illustrates the prototype that has been used for scanning. The camera in this prototype is attached to the cover of the laser source and two centimeters to the back to provide a better view angle due the small view angle of the camera (65 degrees). Therefore this shifting is needed to let the camera capture the maximum amount of the laser ring. By moving the prototype with constant speed, the camera collects the data and saves it to an external computer.

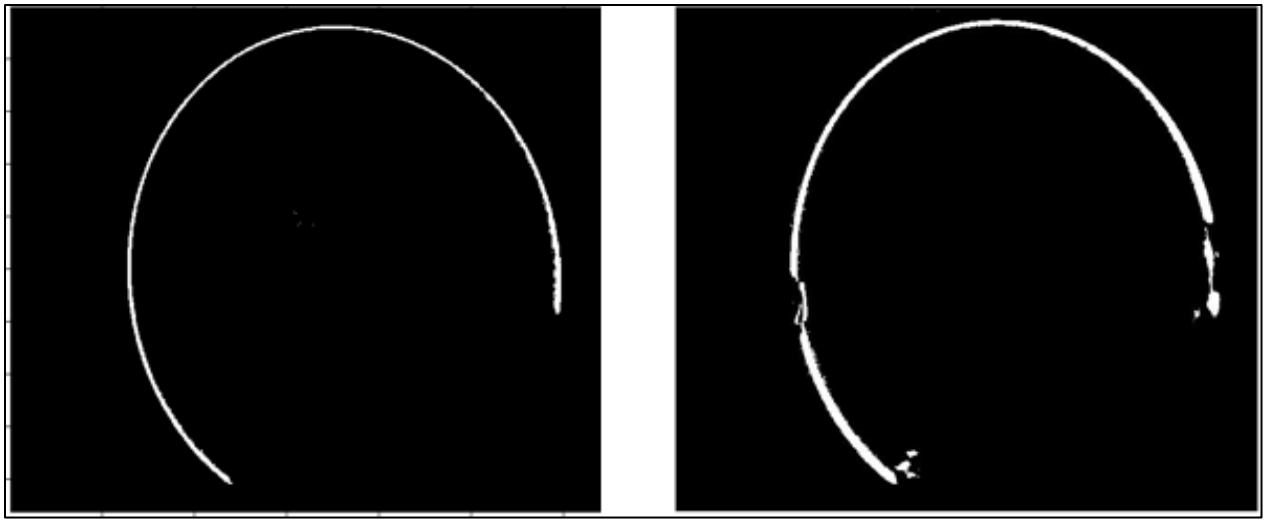
Figure 5-19. Simple camera and source bundle prototype

As shown in **Figure 5-20**, the laser ring is not complete because the camera captures only a partial part of the projected laser ring. Thus, part of the scene is blocked. This kind of missing ring parts is called shadowing. This shadow area blocks about 25% of the whole scene while the other parts can be reconstructed successfully. Therefore, this kind of missing information is one of the limitations that should be considered to solve in the next sections. **Figure 5-21a** shows that the laser ring does not have any deformation. On the other hand, **Figure 5-21b** shows some deformation in the segmented laser ring. The changes in the laser ring shape are indications of the existence of deformation on the pipe's wall which is decoded later to form the damage 3D profile.



(a) (b)

Figure 5-20. Some frames form the collected data: a) with no defects and b) with defect



(a) (b)

Figure 5-21. Results after applying the thresholding

3D image reconstruction

As mentioned before, the reconstruction process depends on the triangulation between the camera and the projector. To determine the shifted location of the projected lines, some intermediate image processing steps are implemented. The complexity of the required steps depends on the type of the used structured light methodology and the amount of noise and artifacts that corrupt the camera image. For this project, two types of implementations are considered. The first one is the single laser ring, and the other is the multi-ring multi-color method which will be implemented in later sections.

For the single laser ring, the reconstruction algorithm is applied to either one of the RGB layers that has the maximum intensity for the laser ring or the gray scale version of the image. In the case of the red laser scanner, the green layer is chosen because it contained minimum amount of noise from the red light. The extraction process starts with applying a weighted threshold to the image to reduce the noise. After that, a two-dimensional Sobel filter is applied to extract the edges the ring from the image. To enhance the laser line detection, a multiple morphological operation is applied to fill and connect the discontinuities of the laser ring [40]. The laser ring is then extracted, and the deformation depth is calculated. **Figure 5-22** shows a brief description of the implemented algorithm.

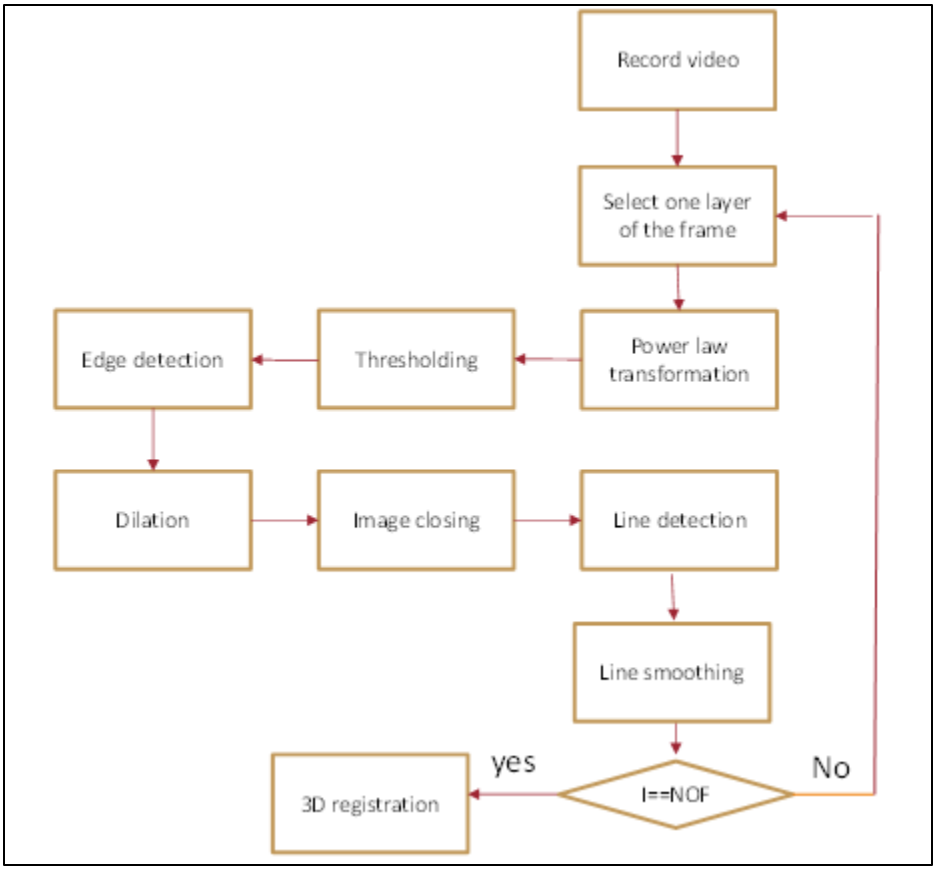


Figure 5-22. Reconstruction algorithm for single laser ring

The Multi rings patterns are proposed to increase the resolution and the speed of the 3D acquisition. With the introduction of multiple rings in the image, the problem of the ring correspondence arises. One of the proposed solutions to solve this problem is the use of colored rings. The colors of the rings are coded with a special code to create sets of unique sequences that can be identified correctly in any location inside the image. Many methods are proposed to decode those patterns. One of the methods is to use graph based reconstruction to decode the pattern [41, 42]. In this method, the image is segmented with watershed transformation. Then a graph diagram is created for the segmented image, and the edges are extracted per this graph diagram. For both described methods, after processing each frame, the depth information is collected and registered to form the final 3D shape. One of the registration methods is performed by using the location of the camera as a reference for the registration process. The other method is to use special image registration algorithm to estimate the location and register the frame. Multiple registration algorithms are proposed to solve this problem, one of the popular algorithms depends on Iterated Closest Point [41]. The problem with those algorithms is that they fail to deal with flat smooth surfaces due to the lack of any distinctive features [42]. For this reason, a combination of both position data and registration algorithms need to be applied to achieve the accurate results.

Preliminary results of the 3D reconstruction for the single laser ring can be seen in Figure 5-23. The results show that about 75 percent of the pipe's surface is reconstructed correctly, and the defect are represented as a bumps or holes on the pipe's walls. The surface reconstruction can be enhanced by using a more advanced setup for the camera and the projector to reduce the amount of the shadows. The results also show some defects due to the instability of the camera holder. A more stable holder and a correction algorithm are required to reduce those defects. The projector and camera setup and the required algorithms for the multi ring methodology are shown in the next sections.

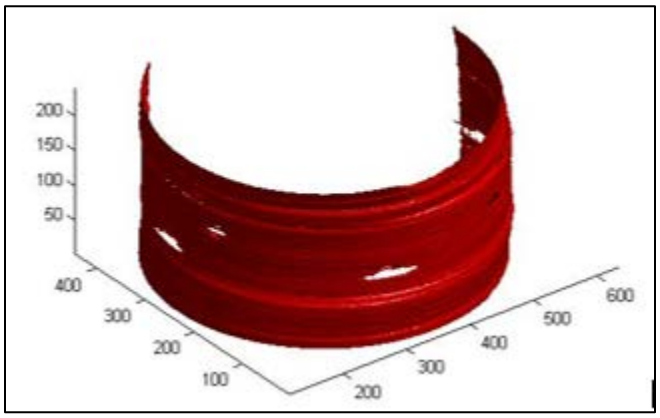


Figure 5-23. 3D reconstruction of the internal surface of the pipe

Scanning with large view angle camera

In this section, we repeated the same 3D reconstruction but with a different prototype. The missing part in the previous 3D image causes some problems. For example, if there are any defects or cracks in the missing area, the 3D reconstructed image will not show these defects. Therefore, some enhancement needed to be made to avoid this kind of problems.
Using a large view angle camera is one of the solutions that make the reconstructed image better. The difference between this camera and the previous one is the view angle. In the old one the view angle is about 65 degrees, but in the new one it is 180 degree that allows the whole scene to be detected by putting the camera and the laser source on the same base line. This wide angle is achieved by using a convex mirror or using a fisheye lens. In our setup, we opted to use a fisheye camera to simplify the setup and make it more robust and practical. **Figure 5-24** shows the fisheye camera that is used in this experimental work.

Figure 5-24. Fish-eye camera

In the previous prototype the camera was attached to the laser source but in the new prototype the camera and the laser source are separated, and they are on the same base line. Because the camera has 180-degree view angles, all the ring will be captured without missing any part. The new prototype can be illustrated in **Figure 5-25**. The prototype will be inserted into the specimen to collect the data. However, the scanning will be performed by human hand so, some misalignment will be introduced. After that, the same 3D reconstruction process will be repeated. Then taking a few frames from the recorded video, the laser light inside the pipe will be as shown in **Figure 5-26** a and b. In **Figure 5-26 a**, the ring is complete and the laser light does not have any deformation. This indicates that the areas that have been covered by this light are still without any defects. On the other hand, **Figure 5-26 b** shows the small deformation in the laser light.

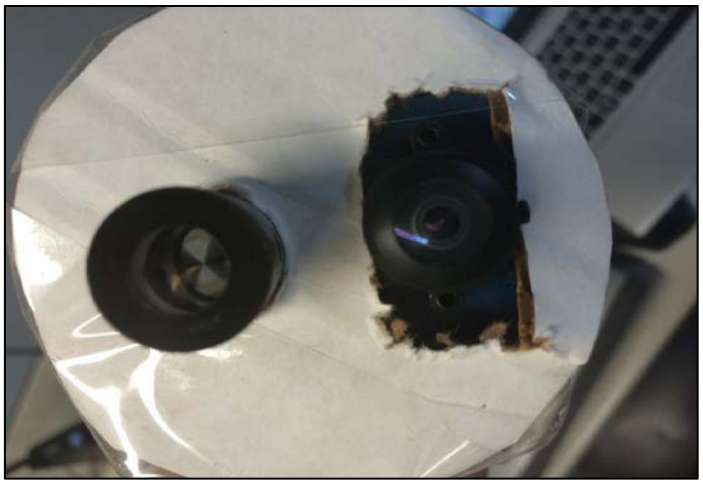
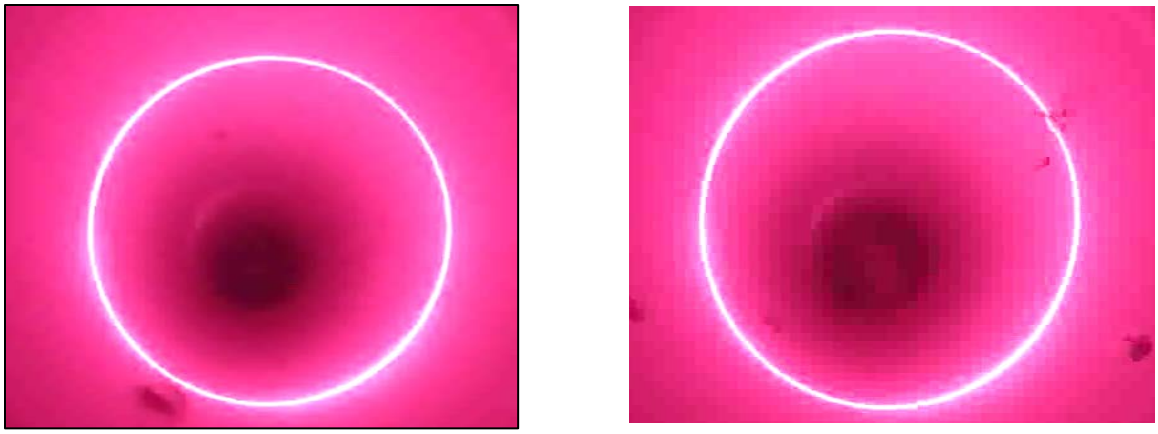


Figure 5-25. The camera and the laser source installed in temporary prototype



(a) (b)

Figure 5-26. Complete ring images, a) full ring. b) ring with some deformation

Per this deformation, this frame will represent the area that has the defect in the pipe. Likewise, by stacking all the frames and generating the 3D image, the defects will clearly appear in the final image **Figure 5-27**. The laser and camera take up a large space inside the pipe making the minimization of the prototype more complicated. Different prototypes are being developed to minimize the size of the sensor and will be reported in next sections.

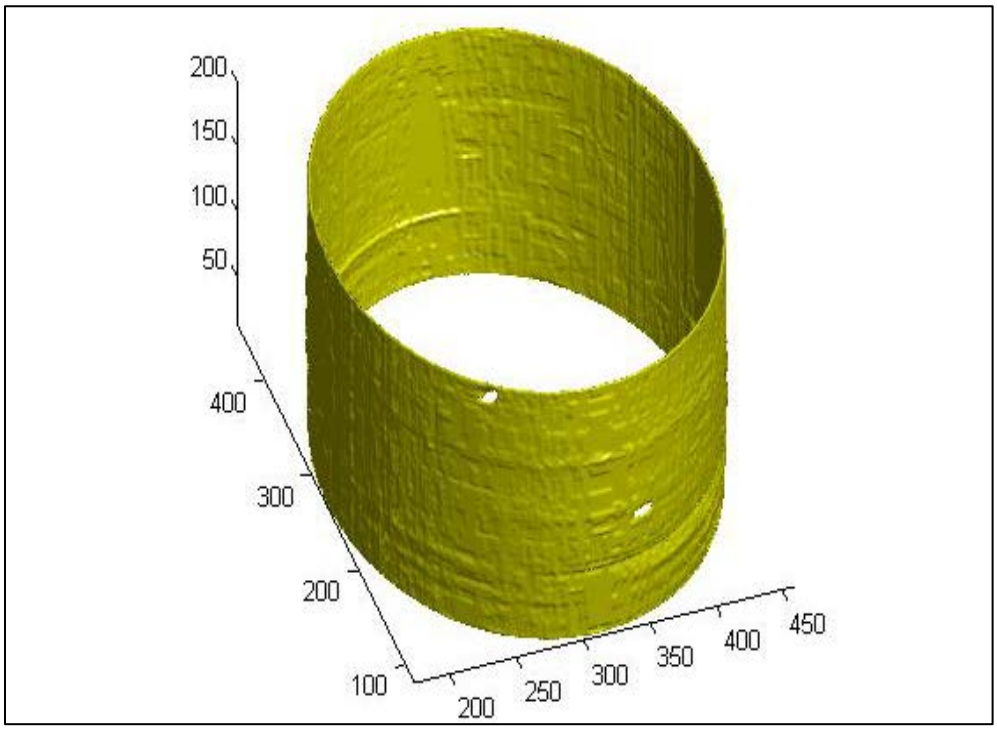


Figure 5-27. The 3D reconstructed image with complete FOV

Scanner and the Sensor in Opposite Direction

As mentioned in the previous section, the laser source and the camera take more space if they are installed in the same baseline. Therefore, in this section different prototype will be illustrated. Although this prototype will reduce the size of the scanner, the image will not be a full ring. Some missing part will be shown in the collected data because of the metal that has been used to connect the laser source and the camera. A part of the light will be blocked and will not be recorded by connecting the prototype as shown in **Figure 5-28**. Therefore, when moving the prototype inside the pipe, the scene will be as shown in **Figure 5-29**.

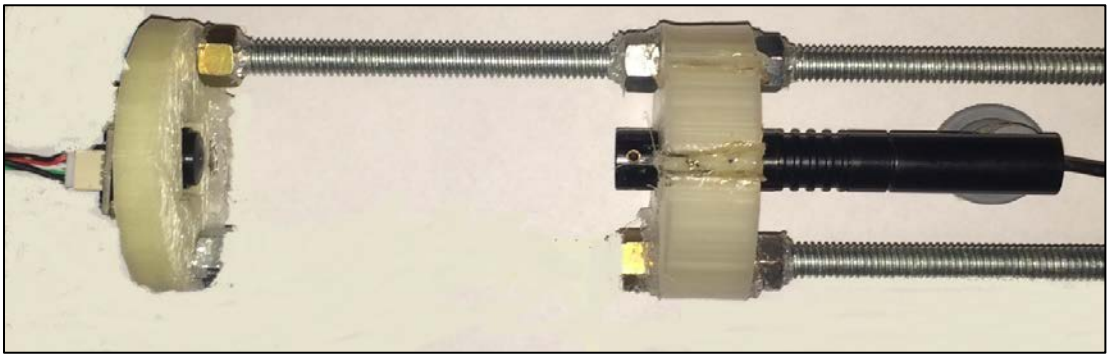


Figure 5-28. The prototype that has been used in this section

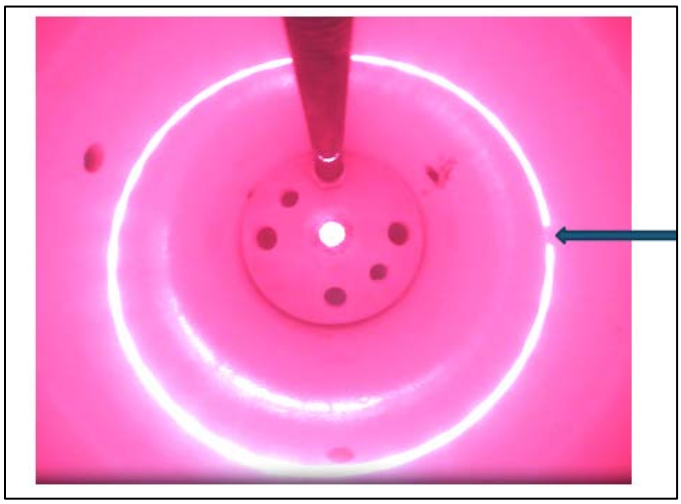


Figure 5-29. Illustration of one frame that has defects

Clearly, the defects can be illustrated as missing points in the laser beam. Thus, after processing all the frames, the final 3D reconstructed image will be as shown in **Figure 5-30.**

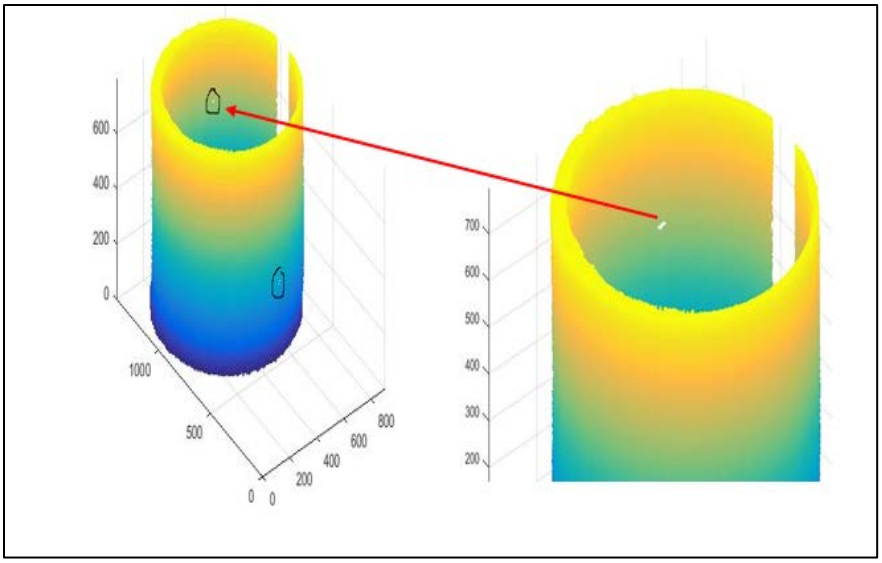


Figure 5-30. 3D reconstruction

Deformation Detection

3 D visualization is an efficient method to emphasize the deformation in the pipes. The circularity of the pipe can be checked by performing eccentricity calculations from the data. In this section, the eccentricity is calculated, quantitative analysis is performed, and the misalignment correction is provided to minimize the camera misalignment.

Eccentricity Calculation

As known, a circle has a fixed diameter or radius in all directions as shown in **Figure 5-31.** This radius is measured from the center of the circle to the edge. Therefore, the distance (a) equals to distance (b). On the other hand, in the case of the ellipse, the (a) and (b) are not equal.

The ellipse has two different radii which are called semi-major and semi-minor. Semi-major denoted as (a) and indicates to the larger radius, and the other radius (semi-minor) denoted as (b). As shown in **Figure 5-32,** there are two focal points which are F1 and F2, and the distance between the focal point and the center is called linear eccentricity. Then by choosing any point on the ellipse (p) and calculating the distance between it and the focal points, we can get the semi-major axes as follows:

$$PF_1 + PF_2 = 2a \qquad \textbf{Equation 5-3}$$

$$f^2 = a^2 - b^2 \qquad \textbf{Equation 5-4}$$

Then by knowing a, b and f, another term can be calculated which is called eccentricity and denoted by e. In case of ellipse, e will be between 0 and 1 (0 < e > 1).

$$e = \frac{f}{a}$$ **Equation 5-5**

$$e = \frac{\sqrt{a^2 - b^2}}{a}$$ **Equation 5-6**

$$e = \sqrt{\frac{a^2 - b^2}{a^2}}$$ **Equation 5-7**

$$e = \sqrt{1 - \left(\frac{b}{a}\right)^2}$$ **Equation 5-8**

For example, if we have a=2.5 and b=2, then $e = 0.6$, Which is less than 1. Also, if a=b=2.5 e, is equal to zero, which is the case of the circle. Thus, by calculating the eccentricity for each frame, we can define if the frame is normal or defective.

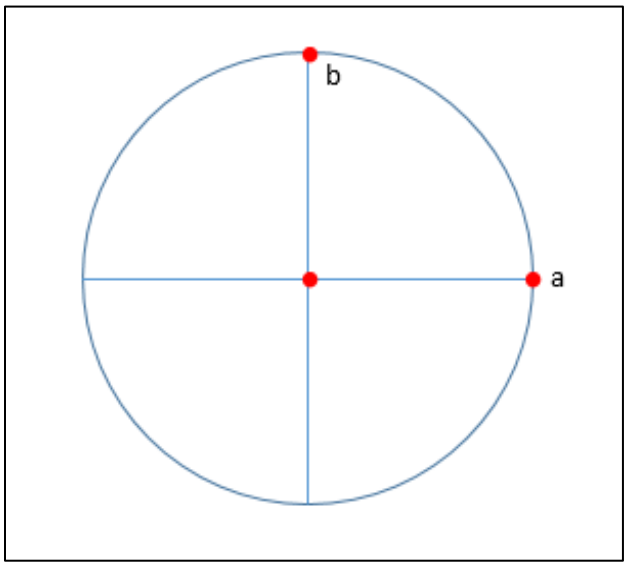


Figure 5-31: Illustration of the circle which has fixed radius

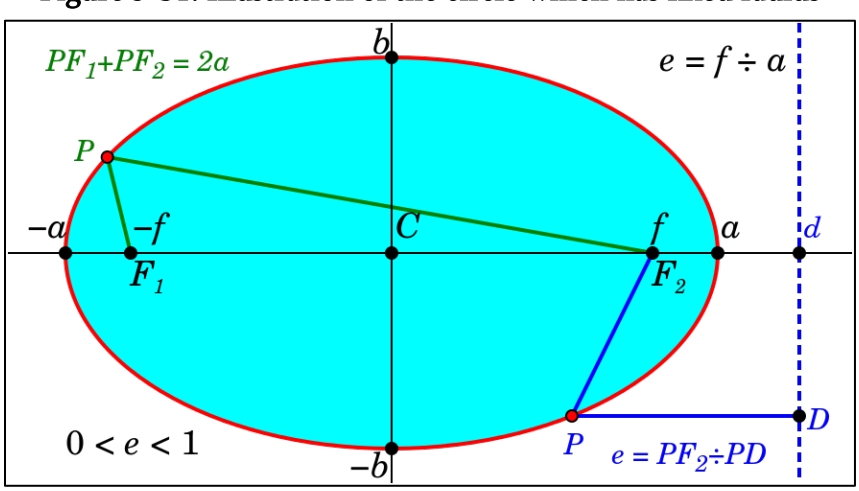


Figure 5-32. Illustration of the ellipse [43]

Eccentricity of a deformation free area

In this section, the eccentricity formula will be applied for some frames that do not have a deformation to emphasize whether if it has a circle shape or not. Because of some other effects, the eccentricity may not be zeros in the case of free deformation area. This issue is due to the scanner misalignment and will be discussed later. However, let's consider that the misalignment is negligible in the current section, and just focus on the large variation of the sequence. Then the defects can be easily illustrated. **Figure 5-33** illustrates the eccentricity of 20 frames without defects, and the variation is not significant. The eccentricity in **Figure 5-33** varies between 0.26 and 0.32. By taking a special case to check the eccentricity. And frame number 60 has been chosen for that, then the semi-major radius is 200 (number of pixels). And semi-minor radius is 193 pixels then the eccentricity will be equal to 0.26622. It is still not zero, but this result will be compared with the results in the following section.

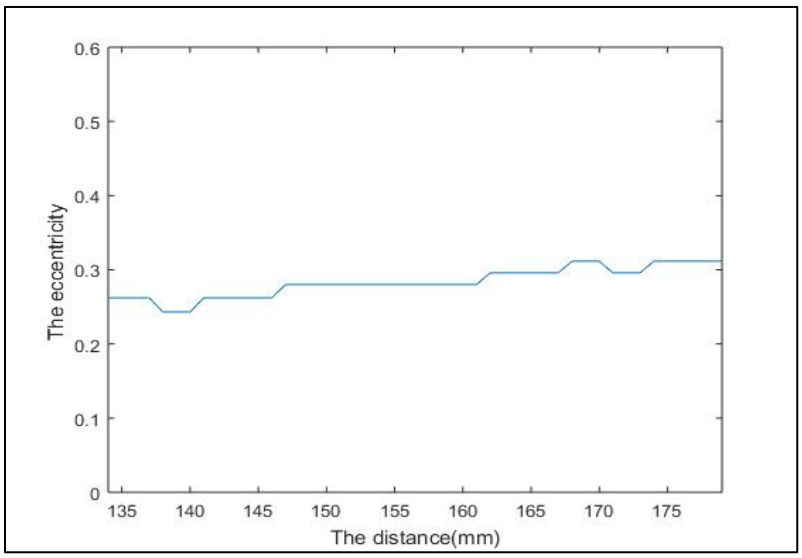
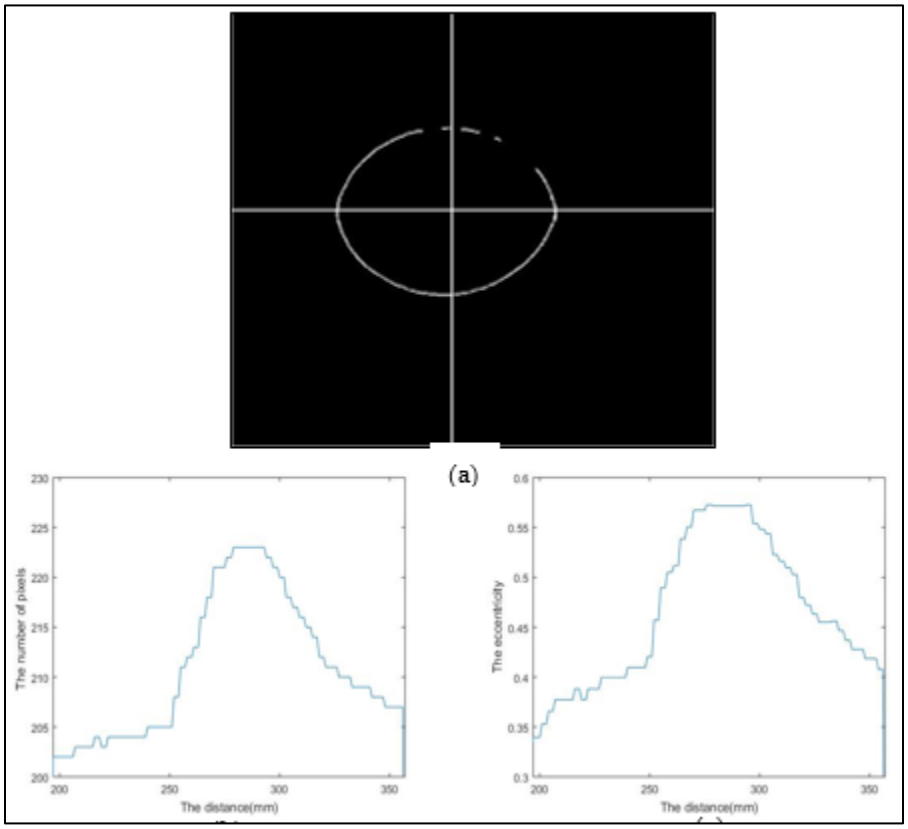


Figure 5-33. Illustration of 20 frames (4.4 cm) from collected data

Eccentricity of the Area with Squeezed Parts

For the frames that have defects, the eccentricity will be much bigger than the other that were illustrated in the previous section. Therefore, by picking some frames that have defects, and let's say ten frames (120-130). One of the frames is shown in **Figure 5-34.a**. Vertical and horizontal lines are added to the image to illustrate the center of the ring. However, the eccentricity will be calculated for frame number 125. Then by calculating semi-major radius which is 217 and the semi-minor radius will be 185 and the eccentricity is equal to 0.5227. The eccentricity is increased when compared with the previous result. To emphasize the difference between the deformation free area and squeezed part the eccentricity will be calculated for all the frames as shown in **Figure 5-34.b**. Initially, the eccentricity will be fluctuating around small value then suddenly will start increasing until reaching the pick

value (0.58) after that it will decrease until starts fluctuating around small value again. These analyses show that the deformation can be easily discovered by using the eccentricity.



b c

Figure 5-34. a) Illustration of one frame with a defect and there are two lines to emphasize the ring center, b) Final result of the eccentricity has been calculated for 71 frames (17.6 cm), c) Illustration of the semi-major radius of all 71 frames (17.6 cm)

Misalignment Correction

The misalignment is one of the problems that are faced when the data is analyzed. This issue makes the reconstructed image not clear enough to emphasize the defects, therefore, in this section, the misalignment correction will be applied to one of the data sets that does not have defects **Figure 5-35**, but the eccentricity is large because of the misalignment.

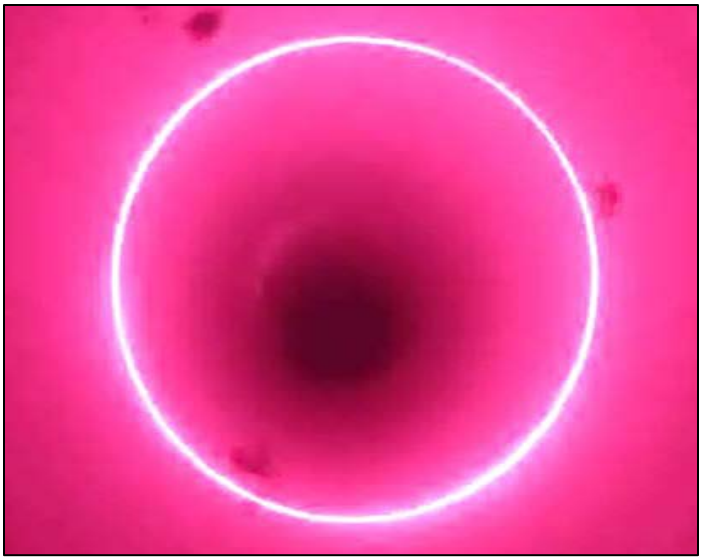


Figure 5-35. The region without defects but the shape is not circle

Moreover, if the eccentricity algorithm was applied to a couple of frames, the output will not be zero as mentioned previously, the eccentricity will be around (0.6359) as shown in **Figure 5-36.** (a). Also, the 3D reconstructed image will be as shown in**Figure 5-36.b**. Thus, the correction should be applied to decrease the eccentricity to be close to zero for deformation free images. This correction depends on the semi-major radius, semi-minor radius, and the orientation of the ring. These parameters are used to determine the (tform) matrix by using an affine transformation, then the image is reshaped and squeezed to a circular shape as shown in Figure 5-37.b.

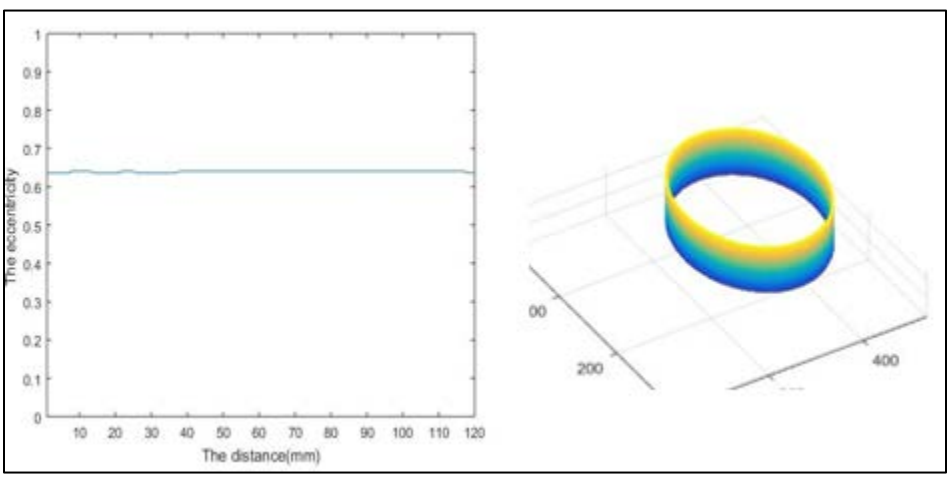


Figure 5-36. a) The eccentricity of multi frames without misalignment correction, b) 3D reconstructed image of number of frames without applying misalignment correction

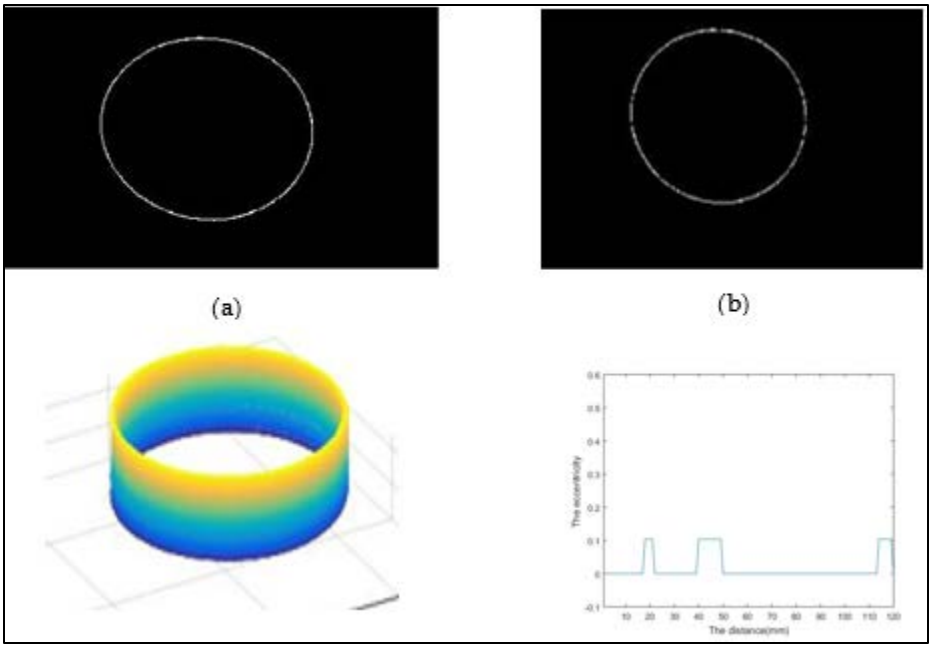


Figure 5-37. a) illustration of one frame that effected by misalignment and the shape is not a circle. b) The same frame after applying misalignment correction and the shape is closer to circle. c) 3D reconstructed image after applying misalignment correction, d) The eccentricity after applying misalignment correction

Finally, the eccentricity of the same frames after applying misalignment correction will be around (zero) which is what we expected to have and Figure 5-37.d illustrates the eccentricity after misalignment correction.

Raspberry Pi based scanner:

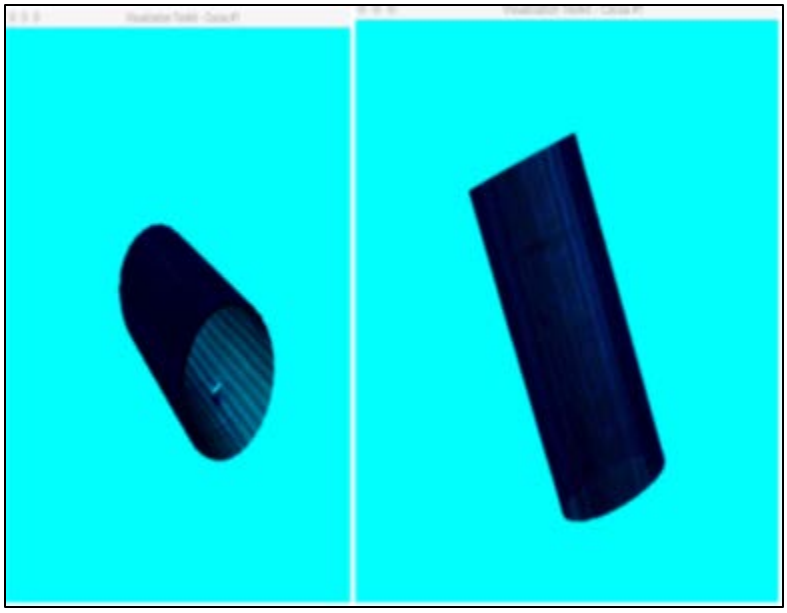
Outdoor field deployment of the scanner makes the mobility of the system a crucial factor to increase its efficiency and flexibility. In this section, a brief description is given about the use of a small size single board computer as the processing unit for the scanning system. This will enable the integration of the system to build a battery-powered handheld device that has a light weight, small form factor, and low power consumption.

The Raspberry Pi 2 model B, was released in February 2015 as the latest Raspberry Pi model with updated hardware. The Raspberry Pi 2 model B, has a quad-core ARM Cortex-A7 CPU processor running at 900MHz with 1024 Megabytes of onboard memory with four USB-ports with one 10/100 Megabit/s Ethernet port, and one micro-SD card slot for storage on board [44]. The Raspberry Pi 2 model B, is faster and has twice the amount of memory compared to its predecessor, the Raspberry Pi Model B+. It now has a quad-core processor, which is speculated to make the Raspberry Pi 2 model B up to six times faster than the previous models. Arch Linux ARM is an ARM-based Linux distribution ported from the x86 based Linux distribution Arch Linux. The Arch Linux philosophy is that users should be in total

control over the operating system, which allows the users to implement it in any way they like. Therefore, Arch Linux can be used for simple tasks and as well as more advanced scenarios. Arch Linux ARM is based directly on Arch Linux, and they share almost all the code, which makes Arch Linux ARM a very fast, Unix-like and flexible Linux distribution. Arch Linux ARM has adopted the rolling-release update function from the x86-version. This means that small iterations are made available to the users as soon as they are ready, instead of the releasing larger updates every few months [44]. The crack detection program is laid out to run on top of an operating system, so installing one on the Raspberry Pi is the first thing to do. The preferred system for this task is NOOBS, as there is a distribution available, which is optimized for the Raspberry Pi. For this project the NOOBS:v:1:7:0:zip distribution was used. The operating system always runs from an SD card, which was found to be reasonable during the development of the Raspberry Pi. SD cards deliver high capacity, are cheap and fast, easily writable and easily changeable in case of damage. The size of the used card should amount to at least 4 GB, as it will contain the operating system (about 2 GB) and should still provide some additional space. We used 32 GB SD card since we were dealing with High-Resolution images. The file system used for the first setup during the integration of the Raspberry Pi into the WSI system was ext4. It is recommended to continue working with ext4 or at least another journaling file system, as journaling provides higher security in case of a power failure or a system crash. After unpacking the downloaded NOOBS image, it can be copied onto the SD card. If connected to a monitor, booting the Raspberry Pi with its new operating system for the first time should display a configuration menu. In the case of not using an extra monitor for the Raspberry Pi the configuration menu can be accessed with the help of the command raspiconfig (as root). One should at least expand the file system on the SD card to be able to use all its space. On top of the operating system, we install OpenCV libraries and Visualization Tool Kit with Python interface. OpenCV libraries provide a set of instructions to perform morphological operations. Visualization Tool Kit installed on this machine helps rendering system for analyzing the cracks. Since Raspberry Pi has a lot of resources available on the system, it provides support for a vast variety of libraries on it. But all the resources available on the system provides only limited processing capability for the current real-time implementation. Onboard memory is very limited about 1GB of RAM due to which we experience a limited Image processing capabilities.

Motion detection and distance estimation:

In this section, the small, low-cost device is used to reconstruct the frames from the camera, with a processing time of 0.35 seconds per frame. The reconstruction results by using raspberry pi are shown below in **Figure 5-38** (a) and (b). As the number of images for processing increases the performance of raspberry pi decreases and the memory available on the raspberry pi is not sufficient.



(a) (b)

Figure 5-38. a and b Represents the Reconstruction Results on Raspberry Pi

To avoid running out of memory, we split the scanning into equal sized slices and store them in individual files. Figure 5-39 shows the slices for certain scanning section. When we split the data, we avoided the memory issue with the raspberry pi, but we were unable to know the distance of scan range. Therefore, to avoid this problem, we calculated the scan distance of each scan sample and also we calculated the distance of defect from the starting point. Since we are moving with the hand, we assumed the scanning speed to be constant and assumed that our sensing module captures ten frames for every 1cm distance. Scan Distance equals to the Number of Frames Captured divided by the Number of Frames Per Centimeter with assumption that the Speed is constant. Then we captured 210 frames, i.e.; we moved 21 cm from the starting point. Then we split the data based on the equal sized slices each of 7cm as in Figure 5-39 (a), (b) and (c).

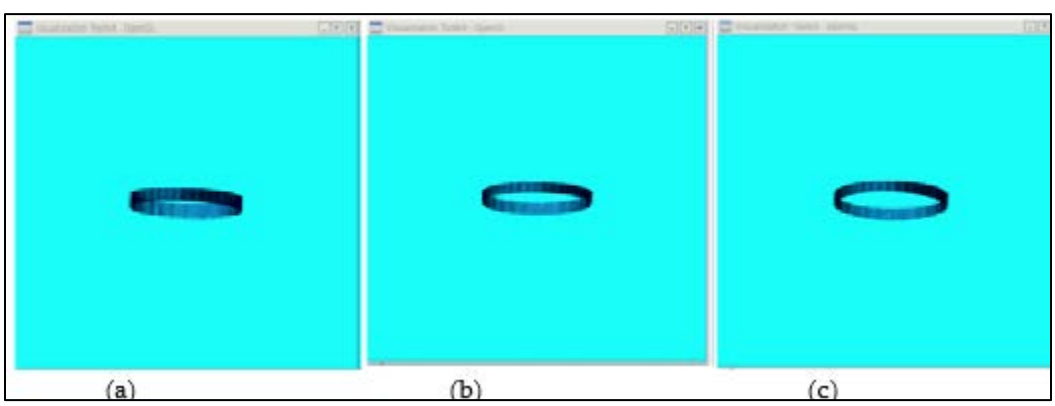


Figure 5-39. Equal Sized Slices 28cm long a) 0-7 cm b) 7-14cm c) 14-21cm

The other experiment was to detect the crack and its position from the starting point. We used our test specimen that is shown **Figure 5-40.** (a) for this experiment and performed some tests. We marked our cracks as shown in **Figure 5-40.** (a). We determined with a measuring scale that the first crack is at a distance of 18 cm, second at a distance of 20 cm and third at 23 cm from the starting position. The rendered results are shown in **Figure 5-40**.(b). The cracks are detected by the principle of motion sensing. At the beginning we capture two images and find the difference between them. If the difference is not null then the motion is detected, and we assume that it is due to some deformation or crack inside the pipe at that point. The distance from the starting point is calculated according to the block diagram in **Figure 5-41.**(a) .

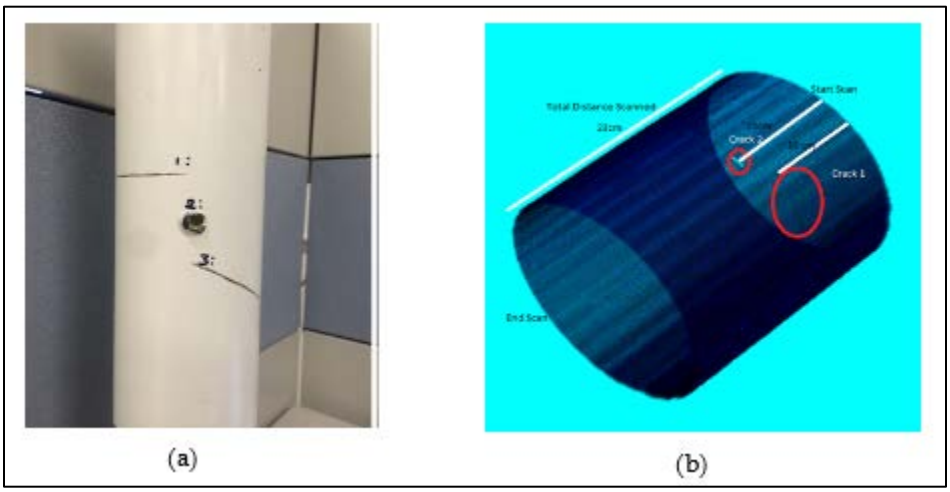


Figure 5-40. a) Object under test for distance calculation, b) 3D Rendered Scan with Crack Detection

Figure 5-41. (b) shows a comparison between the actual and estimated positions of defects. From the figure, we observed that the crack one and two were accurately matching with the original measurements, but few points were not overlapping this was due to the movement of the prototype. Our experiment requires the prototype to be moved at a constant speed manually which is not an easy task.

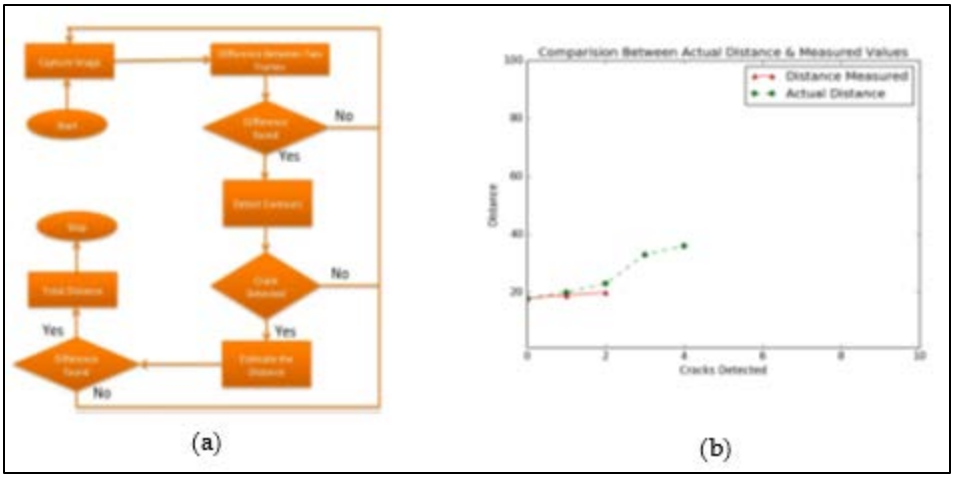


Figure 5-41. a) Experimental flow for distance calibration, b) Comparison between the actual distance and measured distance

Multi-Rings Multi-Colors Scanning

As discussed in previous sections, structured light is one of the methods that have been used to generate a 3D image by using a camera and an active light source. To produce a pattern that is needed to illuminate the scene, a projector is required to do this step. In typical 3D imaging by using structured light, the projector will be like one that is shown in Figure 5-42. However, this kind of projectors will not be sufficient inside the small pipes. An alternative to using a DLP projectors, is design small slide projector to provide the wanted pattern.

Figure 5-42. Structured light with a projector used to generate the [45]

Simulation environment:

A simulation geometry is created by using an open source software (POV-ray) to simulate the structured light scanning process. The simulation helps developing the algorithm in parallel with developing the hardware. It also provides a controlled testing environment to test the robustness of the algorithm against different noise and interference conditions. The

simulation environment provides a light source and a camera that are modifiable to perform different scanning tasks for different targets shapes.

Two geometries were simulated, the first one is a simple rectangular geometry to test the performance of the algorithm then the geometry is modified to simulate the pipe scanner.

1. Simulation with rectangular geometry:

   A 3D object (Stanford bunny [46]) is used as the target to be scanned. In the simulation, the camera and the projector are aligned at different angles with reference to the target. The target is illuminated with a colored sequence that generated with De Bruijn algorithm. As De Bruijn consists of some repeated colors and as the camera cannot differentiate between repeated colors, XOR operator is used to removing any color repetition [30]. A sequence with K=5 and n=3 is shown in **Figure 43.** (a). The image captured by the camera is shown in **Figure 43.** (b), and it is obvious that the projected stripes are deformed by the object surface. After that, the image is segmented as shown in **Figure 43.** (c). The segmentation results show that the area near the tail couldn't be segmented correctly due to the poor illumination because the projector is projecting from the opposite side. The final results after post-processing are shown in **Figure 44**. The results show that the object has been successfully constructed and the depth is retrieved. The 3D acquisition of the object suffers from some small missing parts due to some segmentation problem related to the scene illumination.

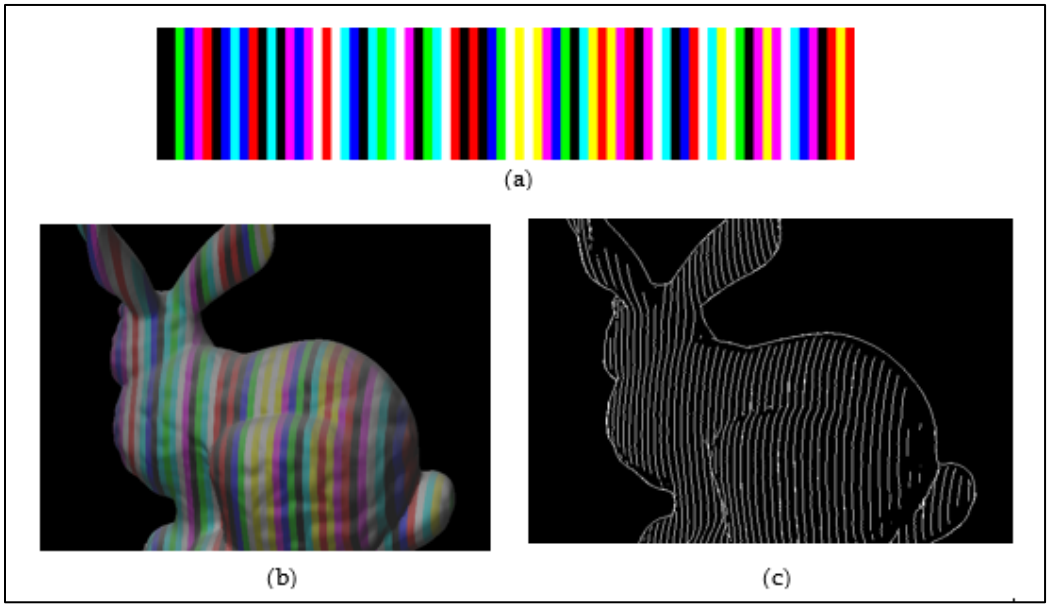


Figure 5-43. a) Illumination pattern, b) Geometry after illumination, c) Segmented image

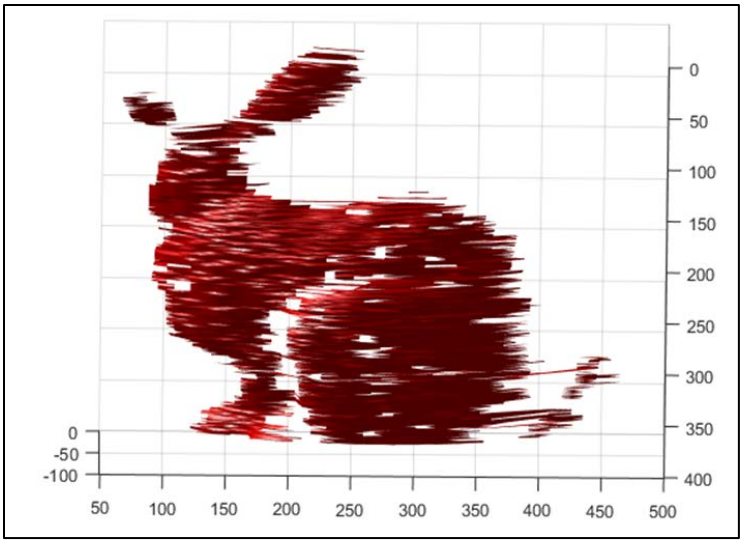


Figure 5-44. 3D reconstruction results

2. Simulation with circular geometry :

The simulation is repeated with circular geometry with the camera and projector orientation shown in **Figure 5-45**. The camera is placed in front of the projector to collect all the reflected light from the pipe walls. The scanner is placed inside a cylindrical tube that was created by using CAD software.

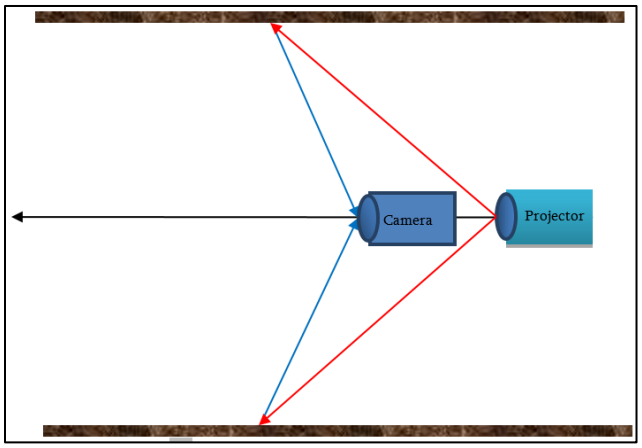


Figure 5-45. Circular geometry of the simulation

The projector illumination pattern is the same as the one shown in **Figure 5-9 b**. It is created by converting the pattern in **Figure 5-9.a** to circular geometry. The first simulated geometry is normal defect free pipe to act as a reference model. The simulation results are shown in **Figure 5-46**. It shows the simulation geometry, simulated pipe surface, reconstructed 3D points from 16 consecutive frames, and interpolated and rendered final profile. The reconstructed 3D geometry shows good agreement with the smooth circular pipe internal surface therefore we decided to move forward and started to introduce defects to the pipe wall to test the algorithm performance. The first geometry is a bent pipe, and it is shown in

Figure 5-47 a. The acquired camera image is shown in **Figure 5-47 b.** It is clear that the projected circular lines are deformed due to the existence of the bend inside the pipe. The reconstructed 3D points are shown **Figure 5-47 c**. The 3D profile of the bend is accurately reconstructed with some missing data on the edges of the bend due to the sudden change in the surface. The interpolated 3D frame is shown in **Figure 5-47 d** in which the missing points have been interpolated. A simpler type of bends is simulated in **Figure 5-48** where it is simulating a flattened upper surface of the pipe due to an external load. The reconstruction is ideal in this case due to the smooth transitions. The second defect geometry type is shown in **Figure 5-49.** It showins the simulation of reconstructing a sudden bump on the pipe wall. From the simulation results it is obvious that the projected rings are deformed by the shape of the bump. But it also shows that the bump is blocking the camera from seeing any detail behind the bump. This is the reason for the partial reconstruction of the bump surface. The algorithm is able to reconstruct on the frontal surface of this bump while the rear part appears as a hole in the 3D reconstruction due to the shadowing effect.

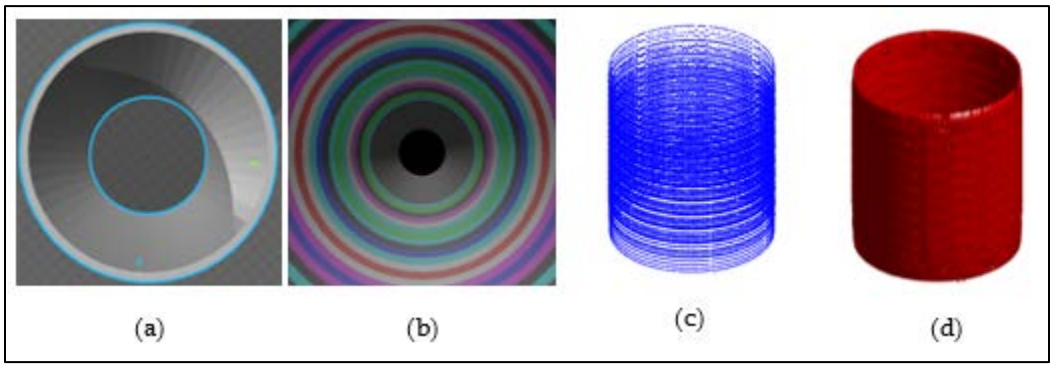


Figure 5-46. Simulation of a normal defect free pipe

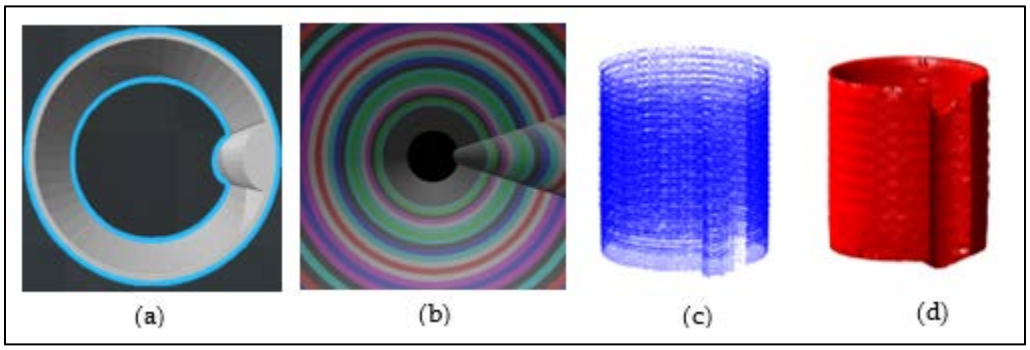


Figure 5-47. Simulation with a bend along the pipe

2. <u>Simulation with a hole</u>:

The 3D sample is modified to have a hole on its internal surface in order to simulate the case of the existence of a hole on the pipe surface. In a similar manner to the previous simulations, the scanner is moved inside the pipe. The simulated geometry and the

reconstruction results are shown in **Figure 5-50.** The simulation starts with the hole on the edge of the projected pattern, where it has no effect on the reconstruction.

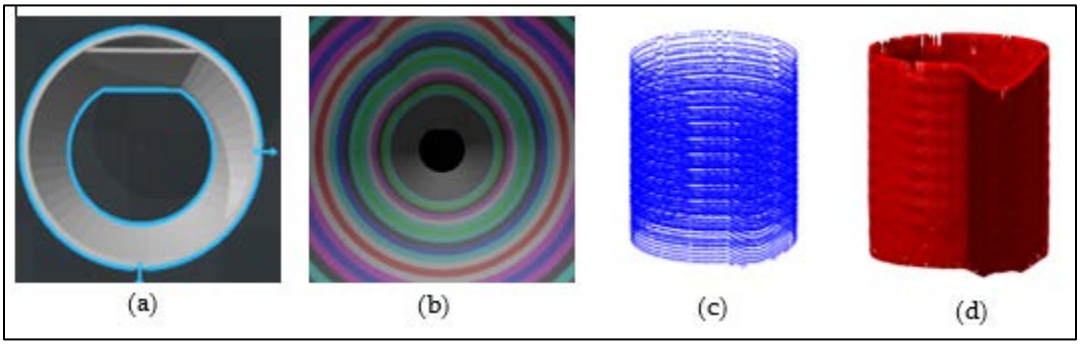


Figure 5-48. Simulation of deformation along the pipe wall

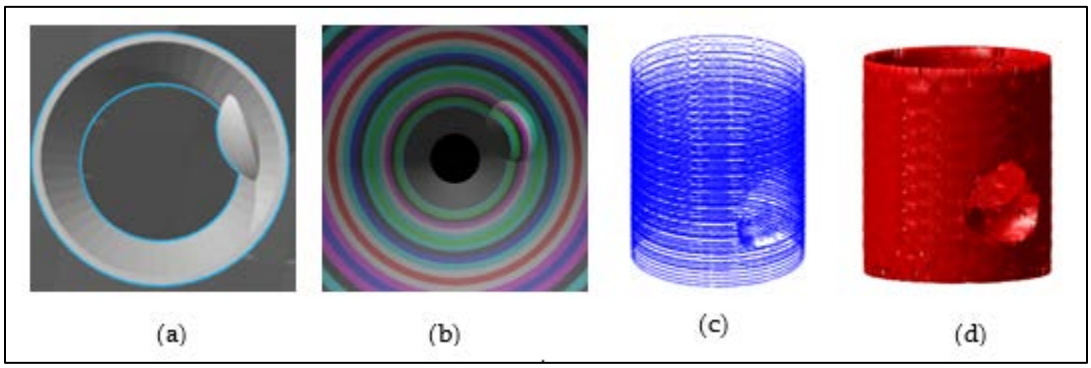


Figure 5-49. Simulation of bump in the pipe wall

After that, it is moved more and more toward the end of the pattern until it exits. The reconstruction results show that at the beginning the hole has no effect, then it starts to create a big hole inside the reconstructed surface (larger than its size) then this effect starts to disappear until it diminishes completely. The reason for this effect is that the hole size is smaller than the size of the projected code. To reduce this effect, shorter codes and thinner stripes are preferred. But for our scanner inside the pipe this effect is not severe, because the final profile is a combination of all the frames in the scanned video. Therefore, the hole boundaries will be reconstructed successfully but with a lower density of 3D points near the hole region.

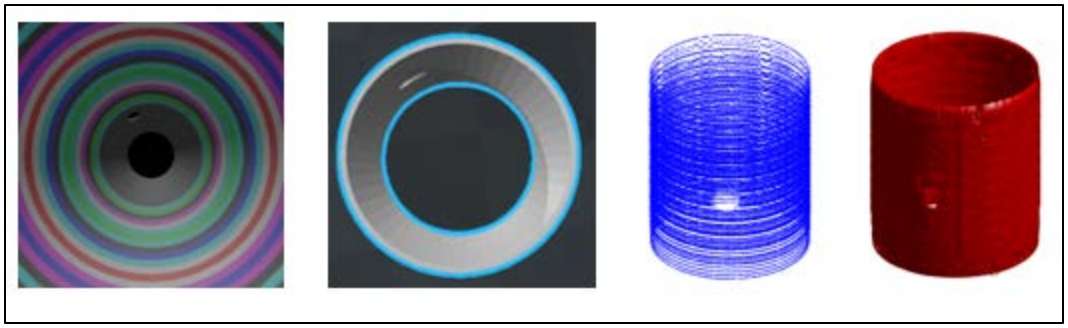
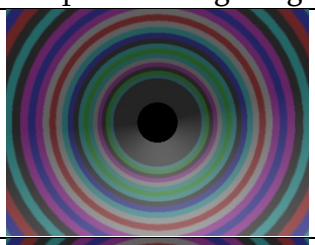
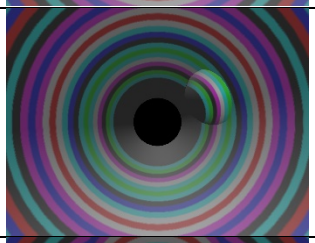
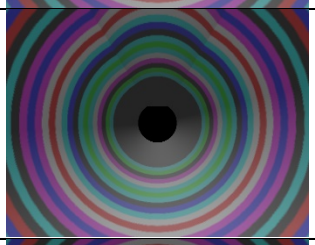
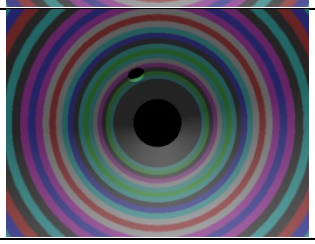


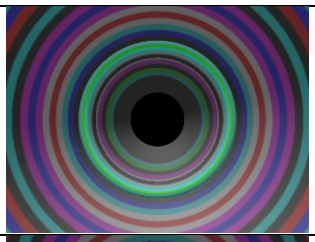
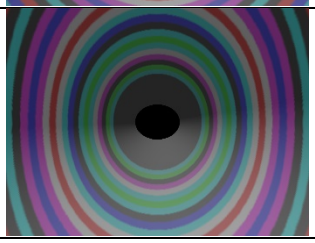
Figure 5-50. Simulation of a hole in the pipe wall

Defects database:

Data reduction and classification tools require a large number of images with different defects types and projected patterns. Therefore, a dataset is created to provide a reference training and validation dataset for our current and future data analysis. The dataset is designed to have a diverse defects types with different view angles and different sizes. The defects types that are simulated in the dataset is explained in **Table 5-1**.

Table 5-1. Database defects types

Damage type	Damage shape	Sample scanning image
No damage		
Impact damage(bump)		
Dent		
Hole		

Split		
Squeeze		

Currently, five defects types have been simulated within the same simulation environment that is described in the simulation environment section, and a defect free pipe is also simulated to take it as a reference. The first defect type simulates an impact damage that is caused by an external tapered object. The second defect simulates a dented pipe by an external load from above. The third defect shows a hole through the pipe wall. The fourth defect represents a split defect on the pipe wall. The fifth defect represents a squeezed pipe by an external load. The images in the data set are created by simulating the scanning process along the pipe which gives different view angles for each different frame in the scanning video. The same defect is then simulated with different sizes to simulate different scenarios. Then the resulted dataset is increased by using image rotation to produce the defects at different locations for each rotation angle. Therefore, the data set will provide a set with different shapes, sizes, location and rotation angles. Noise is added to the dataset to make it more flexible and give the user the ability to control the amount of the noise in the reconstruction or training process. The current data set is a synthetic dataset. Experimental scanning samples will be added later to the give more the depth for the user about experimental effects during the data analysis.

Projector prototype III:

The main idea of using the slide projector is to emit a pattern with different colors of the object to obtain different illumination areas. As a result, a small slide projector has been designed to generate the colors instead of using the regular projector. Moreover, to create this projector, some lenses are needed. The lenses that have been chosen for this step have a diameter (1.5 inches). Therefore, the size of the projector will suite the small diameter pipes that we need to scan. Then, these lenses will be stacked with a certain way inside a prototype that has been designed by using the 3D printer. One achromatic lens, two double convex lenses, and two convex lenses are used as shown in **Figure 5-51**. Every lens has a focal length that should be considered when the designer stacks the lenses together. The main idea of using the Achromatic lens is to reduce effects of chromatic and spherical aberration. The other lenses are used to magnify the image and increase the resolution. The slide that is used in this prototype has multiple rings as illustrates in **Figure 5-52**. Moreover, the power of the light source is strong enough to generate the light that is needed to emit the pattern.

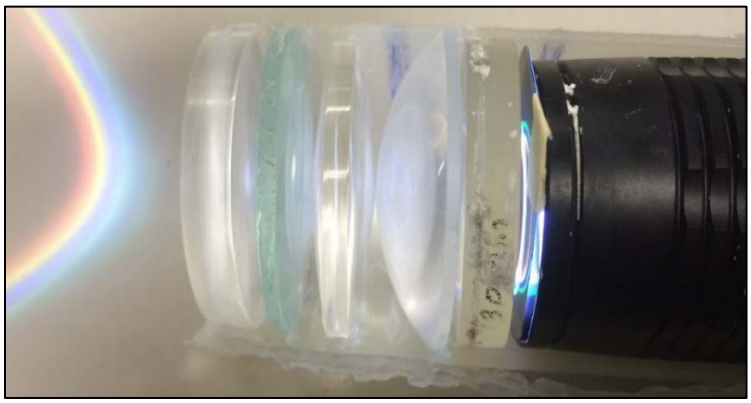
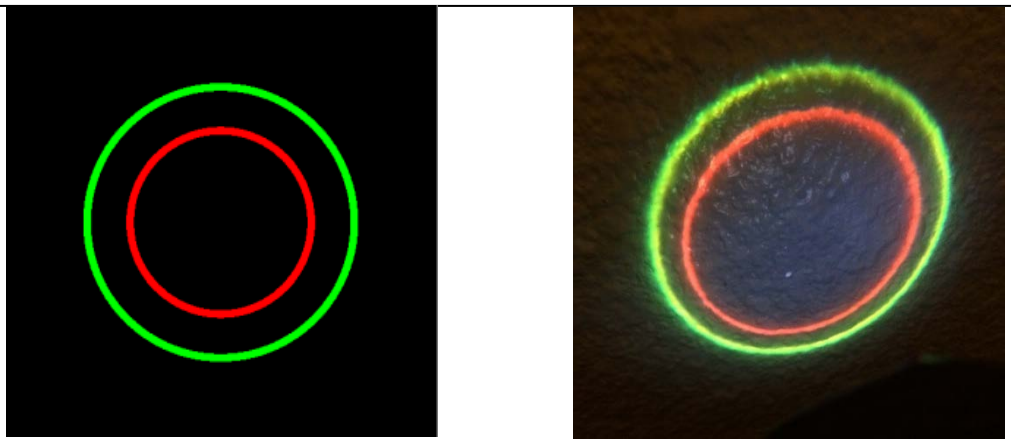


Figure 5-51. The lenses stacked inside the prototype



(a) (b)

Figure 5-52. a) The slide for the projector, b) Projected pattern

However, the slide has two different colors (red and green) as shown in **Figure 5-52.a**. There is a black area between the two rings just to separate the edges. This separation can be removed in the future after increasing the resolution. **Figure 5-52.b** shows the pattern on the rough surface to clarify the edges. The length of the projector is 7.87 inches, and the largest diameter is 1.77 inches which can be used to scan the pipe that has 1.85 inches. Also, the distance between the light source and the rings when are projected inside the pipe is 1.02 inches. Then the next step is to connect the prototype and the camera as shown in **Figure 5-53.**

Figure 5-53. Projector and the camera connected to each other

The camera that has been used in this step is the one that has small view angle. Because the pipe has a small diameter, small view angle camera can be used instead to the large view angle camera. As shown in the figure the distance between the light source and the camera is 5.51 inches. By moving the prototype inside the pipe, the data will be collected by using the camera. Each frame has some features that came from the structure of the pipe. Figure **5-55** (a) illustrates some frames that have been taken from the data. Clearly, the shape of the rings has been changed as shown in **Figure 5-55 (b)**. This change shows the squeezed parts in the pipe. As a result, the reconstructed image will have the same shape of the corrupted image. But in this case, there is some illumination that has been generated from the extra light that comes from the lenses. To block the extra light, a black tape has been used to cover the edge of the lens as shown in **Figure 5-54 (b)**. The difference can be easily illustrated in **Figure 5-56,** and the colors clearer after blocking the white light. Then, the algorithm will process each frame in order to generate the 3D reconstructed image as shown in **Figure 5-57**.

Figure 5-54. a) the projection on the smooth surface. b) the prototype with blocked edges

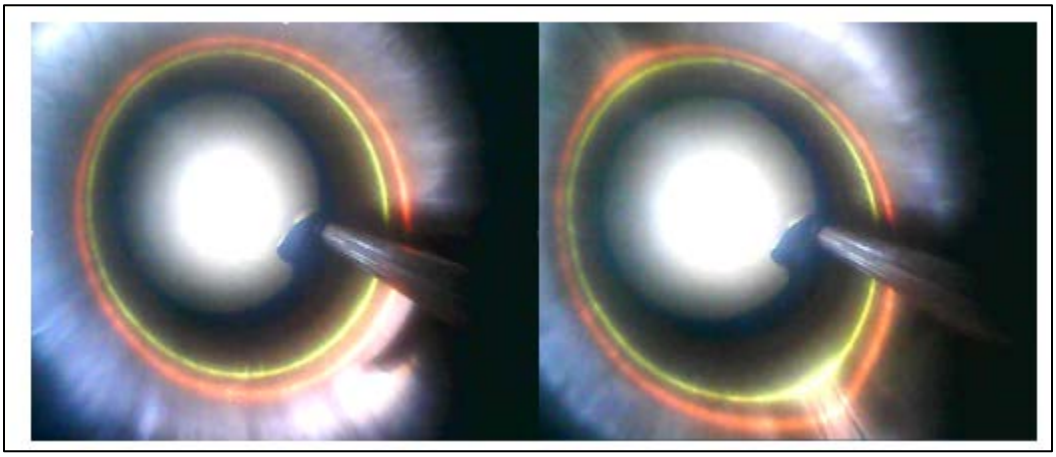


Figure 5-55. a) Illustration of one frame without defects. b) Illustration of a frame with squeezed part

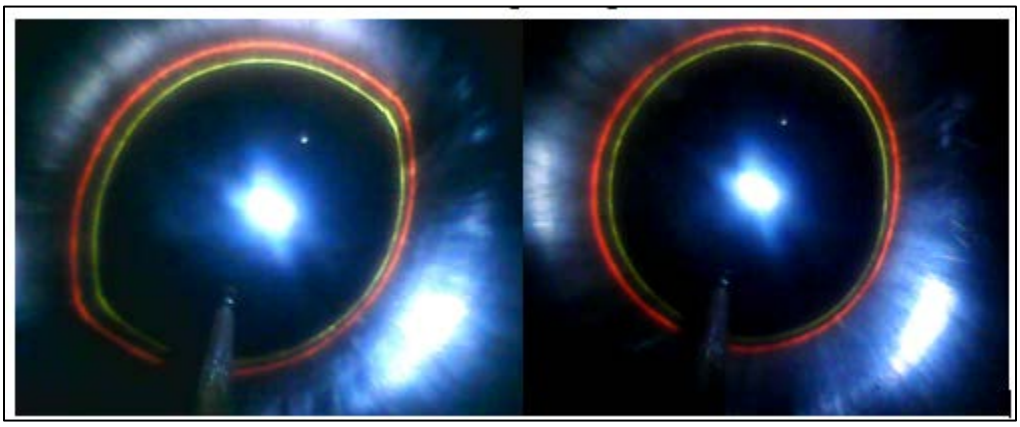


Figure 5-56. Illumination of the white color has been reduced

A color segmentation is done as a first step in the algorithm to generate the 3D image. The red color is segmented and reconstructed as shown in **Figure 5-57**. It can be seen that one of the reconstructed images is not a regular ring shape, but it is twisted. The twisted area in the

image illustrates the squeezed part inside the pipe. Moreover, **Figure 5-58** shows just a few number of frames that show the pressed area.

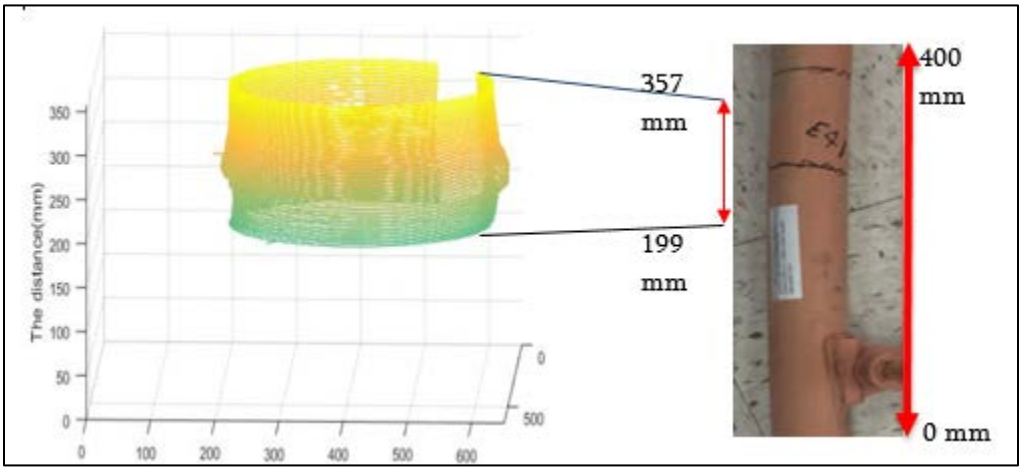


Figure 5-57. 3D reconstructed image with and without the defects.

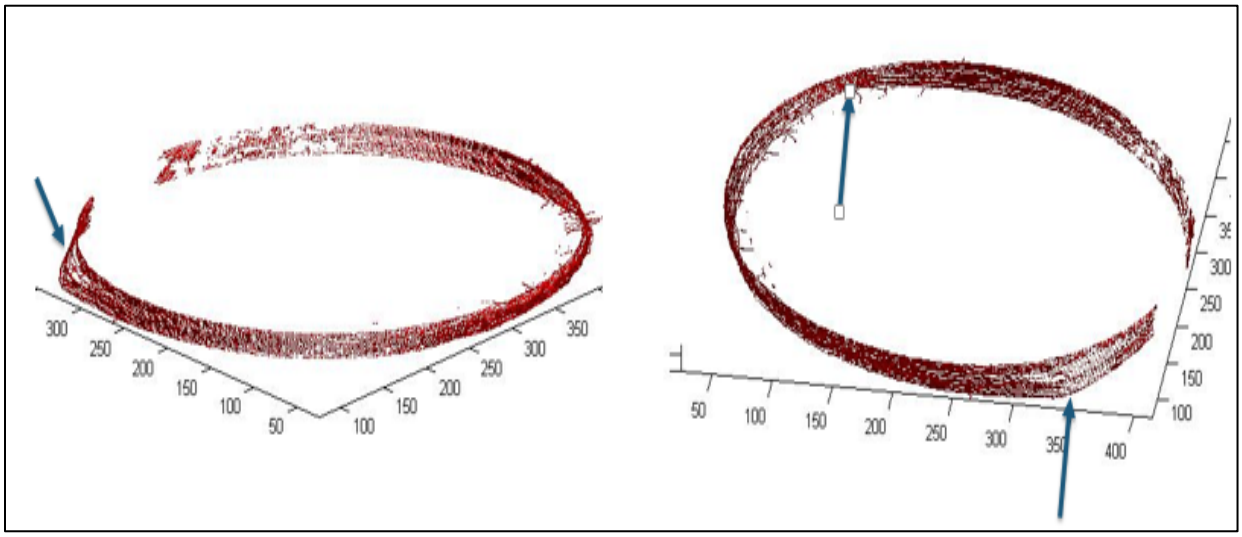


Figure 5-58. Reconstructed frames that show the squeezed parts inside the pipe

Miniaturized sensor (Prototype IV):

Structured light sensors adopt the same stereo concept of using two cameras to extract the 3D information, but use a different path to collect the correspondence points. To solve the problem of finding correspondence points that are found in the passive stereo sensors, the structured light uses an active light source that projects a pattern that has well-defined points. These points can be easily matched when comparing the projected and captured images. The number of reconstructed 3D points is proportional to the number of the matched points between the projected and captured images. Therefore, the design of the

sensor plays an important role in the reconstruction process because enhancing the quality of the projected and captured image increases the number of acquired 3D data proportionally. In this section, we are focusing on designing a new miniaturized sensor that provides a high-resolution projected image, wide projection angle, and a small size to inspect 0.75-inch pipes. The structured light sensor consists of a light module that projects a highly-textured pattern and a camera that captures the deformations in the projected pattern. Three main schematics were considered for the sensor design as shown in **Figure 5-59**. Schematic (a) employs a projector and a camera that are facing each other. In this scheme, both the projector and the camera are required to have a wide view angle to shorten the length of the sensor. In schematic (b), the light module is put in series with a camera that is pointing in the same direction. In this scheme, only the camera is required to have a wider angle than the projector. In schematic (c), the light module and the camera are aligned in a parallel scheme. In this scheme, the camera, and the light source can have similar focal length because both are pointing to the same direction and lies at approximately the same distance from the pipe internal surface.

For a fixed diameter pipe, both designs (a) and (b) could have the same characteristics, but as the diameter of the pipe changes the drawbacks of the first design starts to appear. By increasing the diameter of the pipe, the rays from the camera and projector will no longer intersect at the same area on the pipe internal surface. To solve this problem, the distance between the camera and the light module needs to be increased and thus increase the overall sensor size as well. In schematic (b), the rays from the camera and the projector are propagating in the same direction, which means that the intersection points will only diverge slightly with the increase of the diameter size. The third scheme doesn't suffer from the problem of changing the pipe diameter as both the camera and the light source are aligned in the same direction and is also shorter that the other schemes, but aligning the sensor components beside each other's will increase in the overall sensor diameter. For the reasons above, we opted to go with testing schematic (b) and (c).

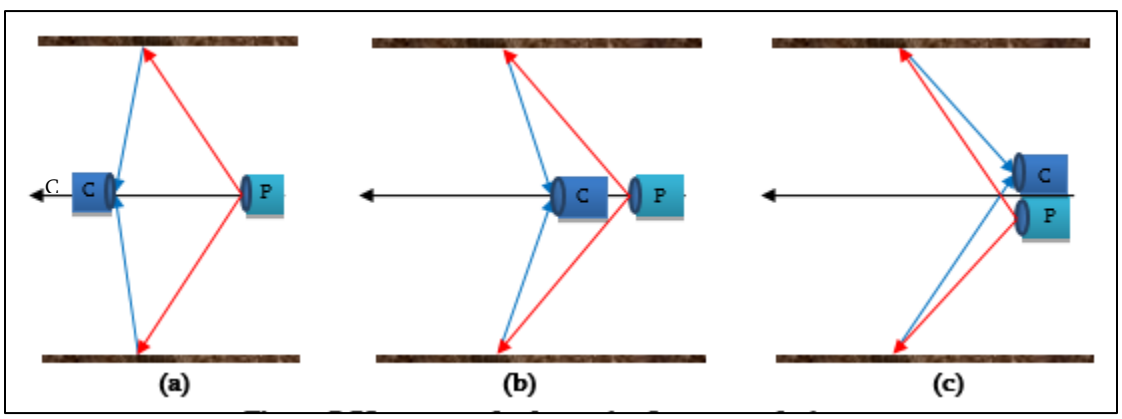


Figure 5-59. proposed schematics for sensor design

As mentioned before the sensor consists of two main components, the camera, and the light source module. The bellow sections give a detailed description of the sensor components, the camera, and the light source module and give a brief description of the fabrication process.

Structured light module:

As the size of the inspected pipes can go down to 1 inch, the diameter of the new light source module is reduced to be less than 0.6 inches. The current digital light projectors in the market are not a good candidate for this task due to their large size (larger than one inch diameter) and the difficulty of customizing these projectors to be embedded in our sensor. Our team chose to go with designing a new small slide projector light module. This type of projectors can be scaled to be in the range of 4 millimeters in diameters [47], and can also provide high flexibility in changing the optical properties of the projector.
This type of projectors adopts powerful white light source to illuminate a colored slide to produce the required image. The resulted image from then projected by using projection optics. The overall structure of the structured light module is shown in **Figure 5-60**. It consists of the following parts

1. Main light source: The structured light module requires a strong light source to produce high contrast images and it should also have a small form factor to reduce the sensor size. There are two main options for this task, either using an external light source and transfer the light to the projector through an optical fiber or using a high-quality light emitting diode. Each one of the options has its strengths and weaknesses. The external light source provides a high amount of power but adds more complexity to the system with an additional external unit and an optical fiber for light transmission. On the other hand, LEDs provide simple and easy to implement a solution but produce a high amount of heat in the sensor head when high illumination is required. Our current prototype is tested with LED light source (Cree XHP50) which can provide a maximum of 149 lm/W while having a maximum output light intensity of 2546 lm under ideal conditions. To enhance the heat convection from the module and to reduce the LED teperature, a high-quality copper heat sink is attached to the LED while maintaining low current at 100 mA.

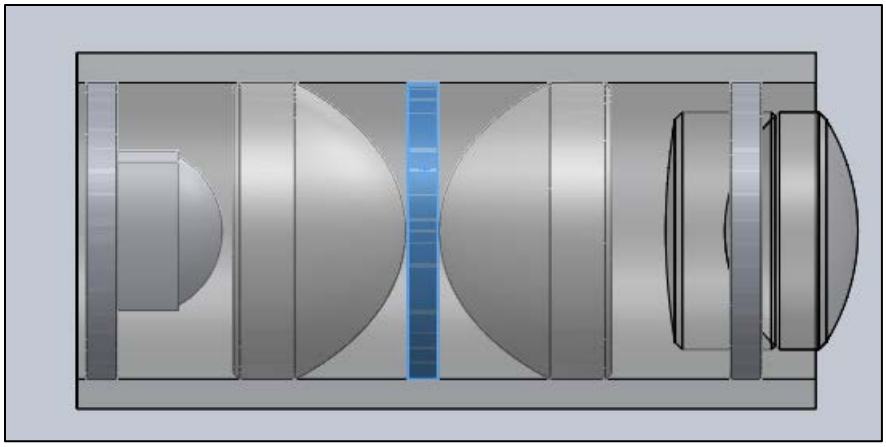


Figure 5-60. 3D schematic of the light module

Condenser lenses: Two condenser lenses with a diameter of 10 mm and a focal length if 8mm is used in the light module. The lens is shown in **Figure 5-62.a**. The first lens is placed directly after the light source to collimate the diverging light rays from the LED and focus them on the colored slide. The second lens is placed before the objective lens to increase the projection angle of the light module. **Figure 5-61** shows the effect of adding the second condenser lens on the view angle of the projector.

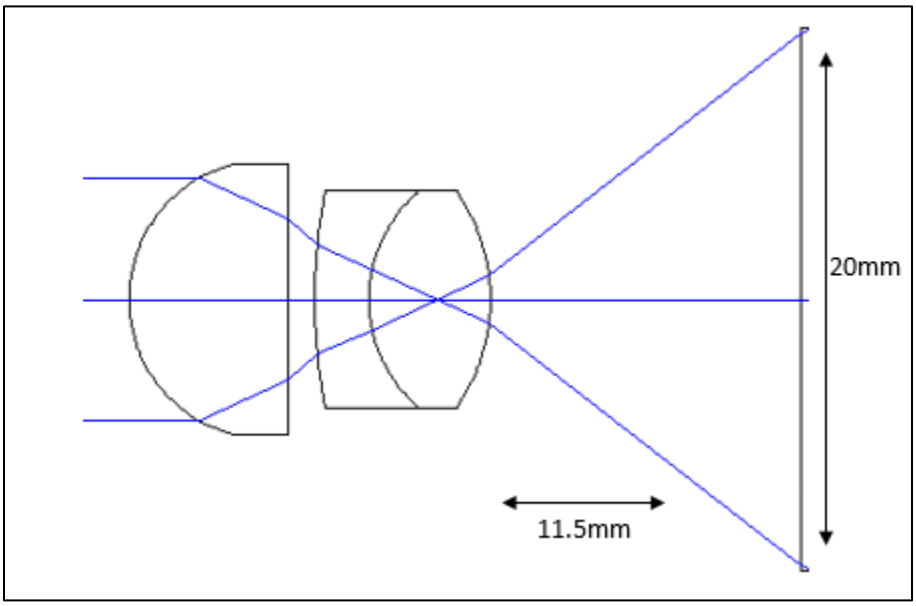


Figure 5-61. Ray tracing of the lenses

2. Objective lens: The image that is produced by the photo slide is passed through the objective lens to project it to the pipe internal surface. An Achromatic doublet that is shown in Figure 5-62.b is used instead of the normal biconvex lens to reduce the

amount of chromatic and spherical aberration. The doublet lens has a diameter of 8 mm and a focal length if 10mm.

3. Photo slide: The photo slide acts as a filter that passes the colors that are printed on the film surface and block all other light components. The photo slide is currently fabricated by using heat resistive transparency film and printed with an inkjet printer. A glass photo filter is considered for future prototypes.

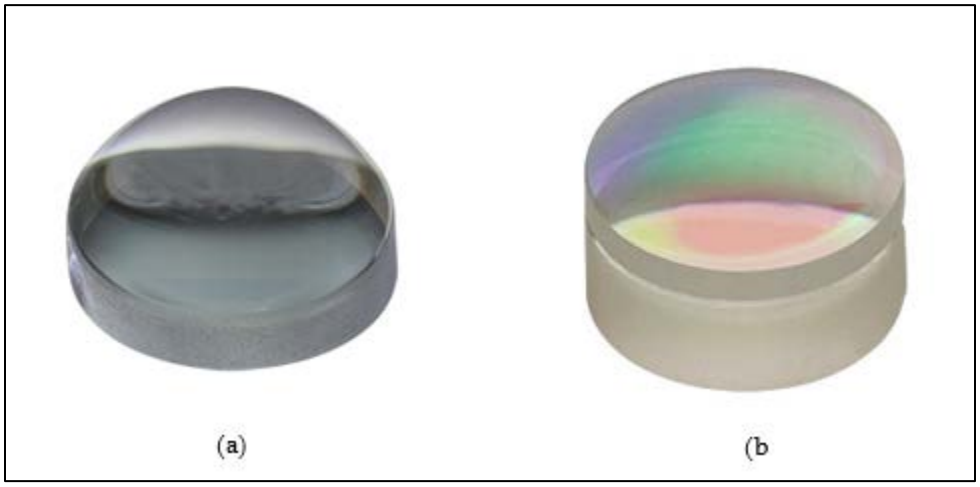


Figure 5-62. a) Condenser lens, b) Objective lens

Camera:

For the camera, we are using small size(10*10*15mm) fish eye camera. This is an analog camera that is connected to the computer by using an analog to digital convertor. The final assembly of the camera and the light module for schematic b is shown in Figure 5-63. The camera and the light module are assembled by using a transparent glass holder to allow the transmission of light through the holder. Schematic c does not require a special component and can be setup directly with the simple fixture to hold the camera and the light module together in parallel.

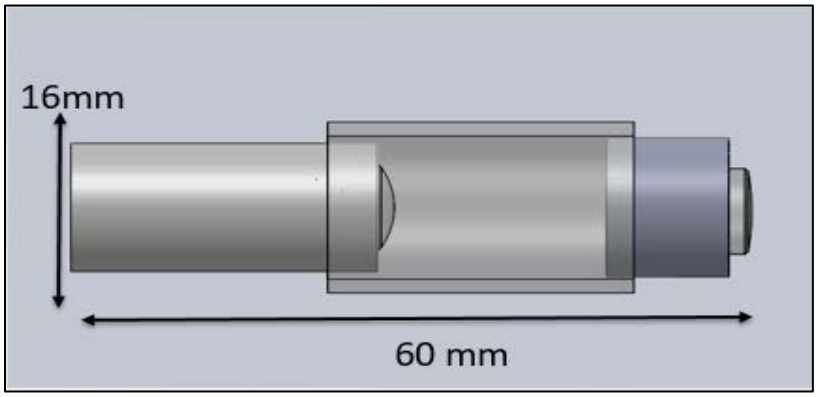


Figure 5-63. Assembly of camera and light module

Triangulation in the scanner

The triangulation process in the geometry at a specific angle θ, is explained in Figure 5-64. $\boldsymbol{c}$ is the camera, p is the projector, P is the intersection point, f is the focal length, d is the distance between the projector and camera, and r is the position of the point on the image plane. For a general camera model, it is defined by the following two equations

$$\frac{f_p}{Z} = \frac{r_p}{X}, \qquad \frac{f_c}{z-d} = \frac{r_c}{X} \qquad \textbf{Equation 5-9}$$

By combining the two equations, Z and X are given by:

$$Z = \frac{d}{1 - \frac{f_c\, r_p}{f_p r_c}} \quad , \quad X = \frac{r_p\, Z}{f_p} \qquad \textbf{Equation 5-10}$$

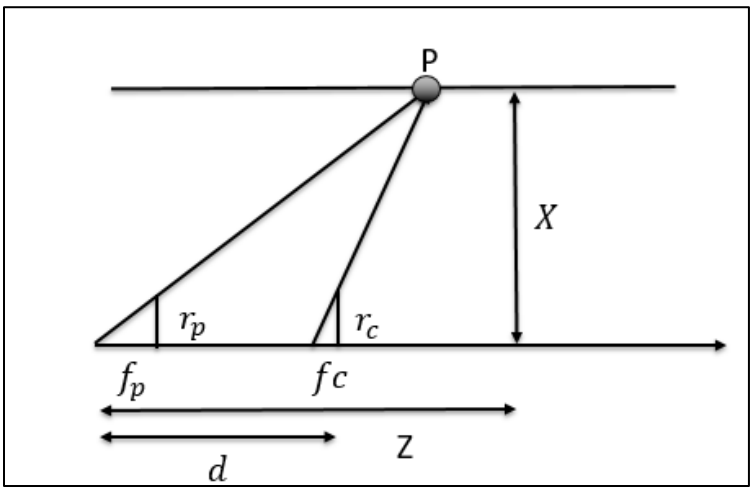


Figure 5-64. Triangulation of the system inside the pipe

Camera Calibration

To reduce the effect of using a fisheye camera and in order to calculate the internal camera parameter for the use in the triangulation process, camera calibration is performed. In this process, a simplified 2D camera calibration scheme that was suggested by [48] is used to calibrate the camera with a 2D checkerboard.

$$S \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ r_{31} & r_{32} & t_3 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$\boldsymbol{u}$, $\boldsymbol{v}$ are the coordinates of the corresponding image point, $\boldsymbol{x}, \boldsymbol{y}$ are the world coordinate of the points, $\boldsymbol{r_{ij}}$ are the rotation parameters for the checkerboard image, $\boldsymbol{t_i}$ are the translation

parameters for the checkerboard image. ***S*** is a scaling factor. The focal length of the camera is calculated and the lens parameters are then used to correct the effect of the wide lens distortion to the image. The fish eye distortion cause the straight lines in the image to be bent. This effect can be removed by calculating the original lens parameters then use an inverse transformation to remove this effects. An explanation of the effect is shown in **Figure 5-65** where the lines of the checker board is straightened after applying the inverse transformation.

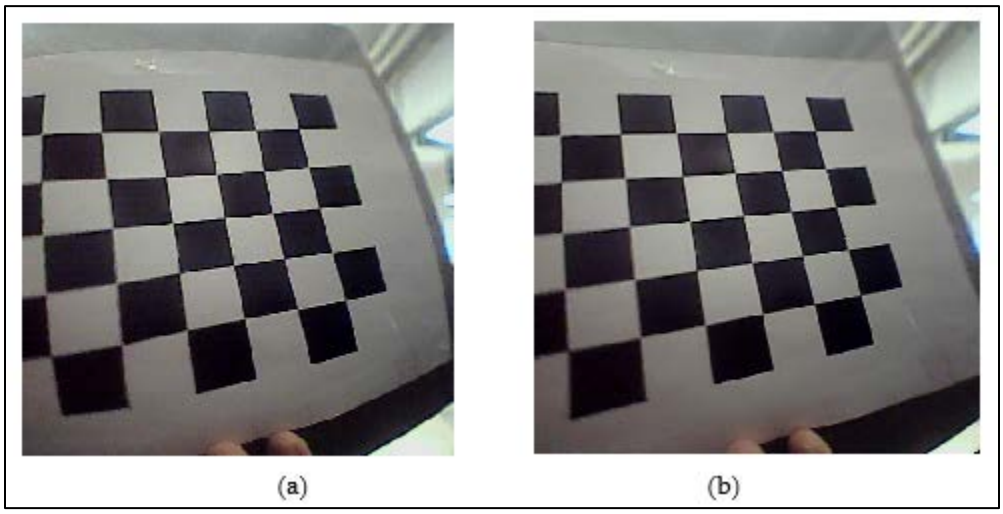


Figure 5-65. Lens distortion removal, a) Before, b) After

Experimental results

In this section, we will explain the structured light module. The integration of the camera will be explained in the next section. To fabricate the light module cover, we are using 3D printing for rapid fabrication and testing before building the final prototype. The cover was printed as two pieces that can be combined to form the final lens assembly. **Figure 5-66** shows one side of the cover. The cover is designed to have 1mm thickness which means that the final light module design will have a diameter of 12 mm (13mm for the final 3D assembled prototype).

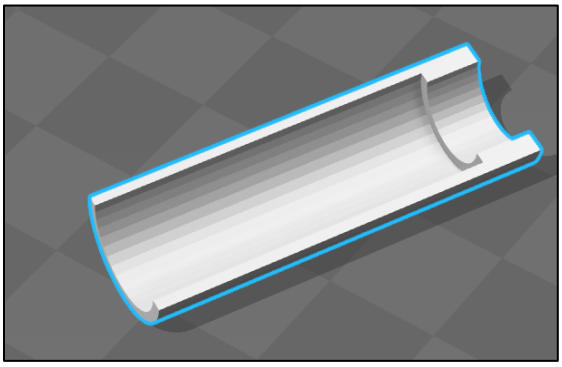


Figure 5-66. One side of the light module cover

The assembly of the components is shown in **Figure 5-67**. Currently, the projector components are assembled and fixed inside the cover by using a glue gun. The glue is used at this stage to facilitate the modification of internal components location in the testing stage. In the future version, the components will be fixed with mechanical fixture after finishing the calibration of the light module. The shape of the final fabricated prototype is shown

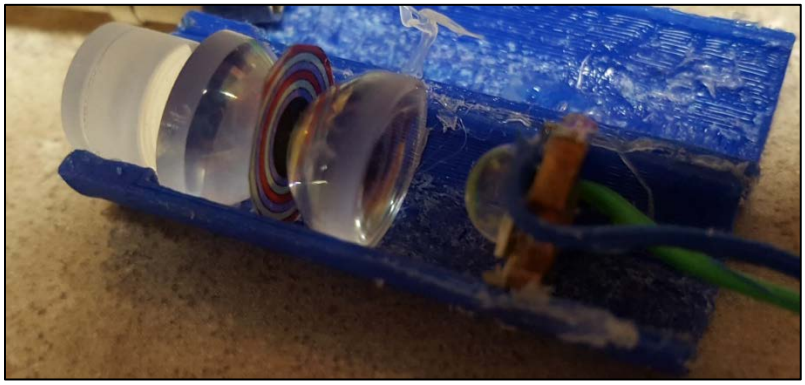


Figure 5-67.Light module internal components

The LED source is excited with a voltage of 5.28 volts and a current of 100milliamps. Different slides were tested to evaluate the performance of the light module. In the beginning, a slide that has four colors (blue, yellow, red, blue) is placed in the light module. The projected colored pattern is shown in **Figure 5-68.**b. The projected image is displayed with high resolution, and color contrast and the edges between the different rings can be identified correctly. After that, we tried to implement a larger number of colored rings to examine the maximum resolution of the projector, but we were limited by the resolution of the printer. The quality of the printer also affects the saturation of the projected colors; the blue rings are projected with purple color.

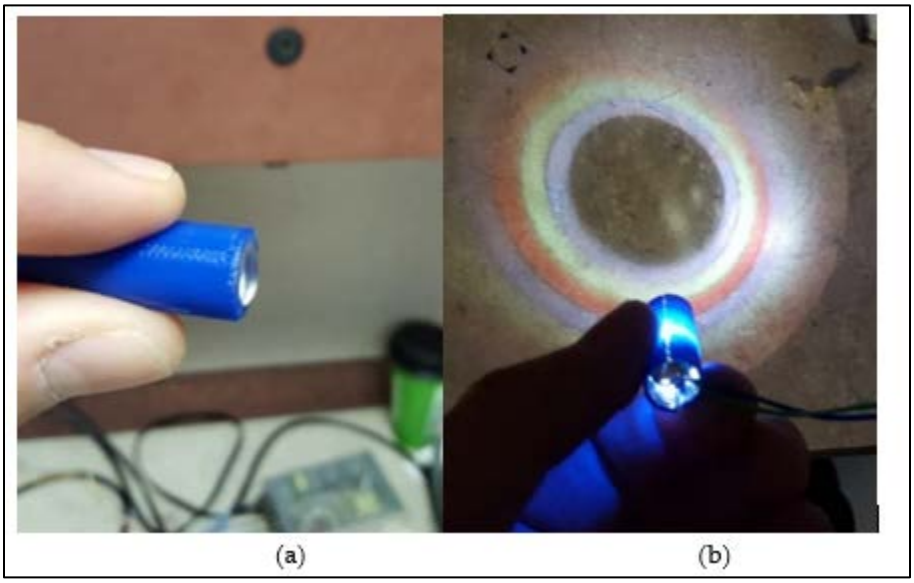


Figure 5-68. a) Projector final assembly, b) Projection of multiple rings

In the design of the final pattern that is used in the scanning, only a combination of red, green, blue is used. This step facilitates the image segmentation on the reconstruction side by providing an image that can be segmented by using the hue channel only [47]. For this type of patterns, the segmentation is performed as follows. In the first step, the image is converted from the RGB to HSV space and only the hue channel is extracted. In the second step, a median filter is applied to reduce the noise from the camera sensor while preserving the edges of the rings. In the final step, a multi-level thresholding is performed to isolate each different color; then the edges are extracted by taking the image gradient. The final assembly of the structured light module with the camera is shown in **Figure 5-69.** In this prototype, the slide projector is integrated with the wide-angle camera and connected with a transparent glass tube.

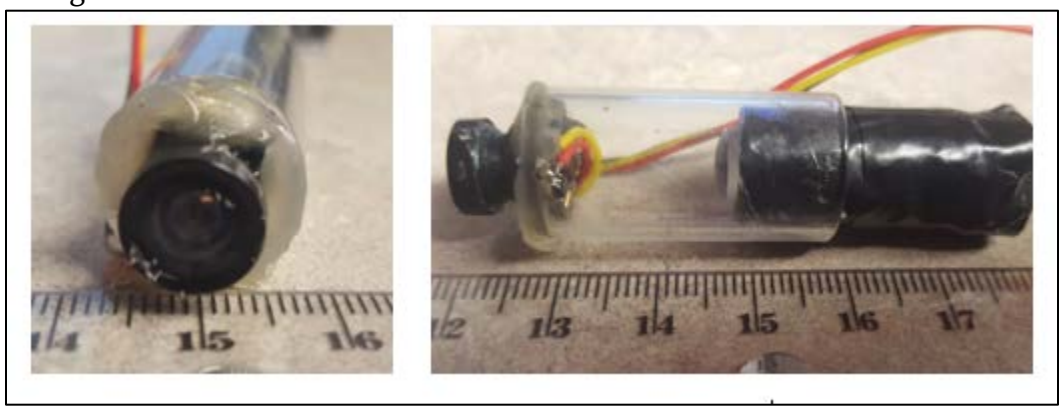


Figure 5-69. Scanner final assembly

In order to perform a controlled scans with a constant speed and orientation, the scanner is mounted into a horizontal scanner as shown in **Figure 5-70 a**, and a single frame of the recorded video is shown in **Figure 5-70 b**. The recorded frame is showing three complete colored rings with the effect of the connection wires appearing as a dark shadow on the upper side of the image. The scanned object is a black PVC tube with a surface texture that is appearing as horizontal lines along the pipe surface.

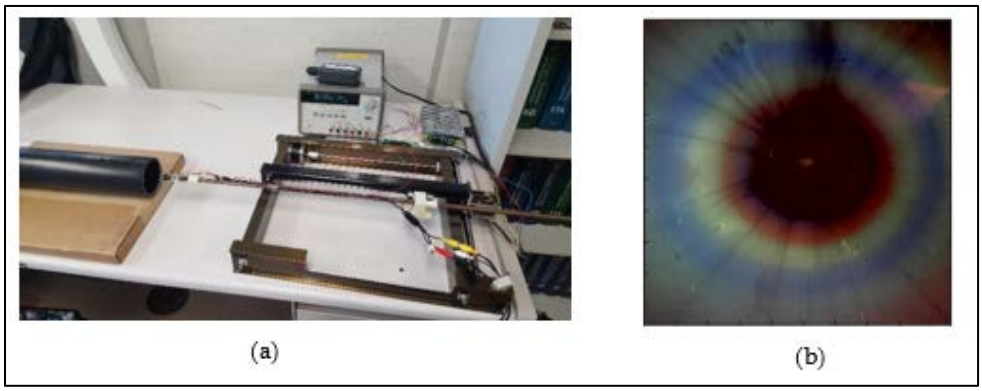


Figure 5-70. a) Prototype IV scanning system, b) a single frame from the system

The segmentation is performed on the image, and the location of each color is decoded by matching it with the color change in the projected sequence as explained chapter one. The conversion from RGB to HSV highly reduces the effect the surface texture and eliminates its effect on the final reconstruction. The results of decoding one frame are shown in **Figure 5-71.**

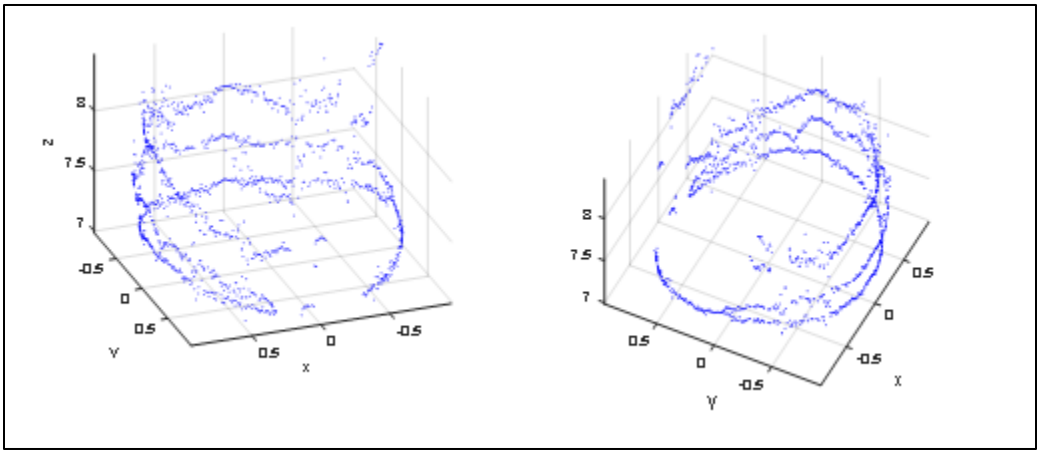


Figure 5-71. Reconstruction of a single frame

The final reconstruction of this single frame is showing the results from the three color transitions. The circular shape of the pipe is reconstructed successfully although the algorithm has failed to reconstruct the complete set of edges from the second and third projected circle. Currently, one of the factors that are affecting the reconstruction process is the spatial resolution of the projected rings on the walls of the pipe. The registration of the consecutive frames is another factor that's is under study because currently, the scanner needs to be moved at a constant speed to assume that the distance between each two consecutive frames is constant. An interpolated result of stacking all the consecutive frames while assuming equidistance between them (which is partially true in our case) is shown in **Figure 5-72.** The current results show that the profile of the scanned pipe is reconstructed successfully but there are holes in the reconstructed profile and there is shifting between the consecutive frames due to the shaking in the platform. Another factor is the existence of outliers from the wrongly decoded edges.

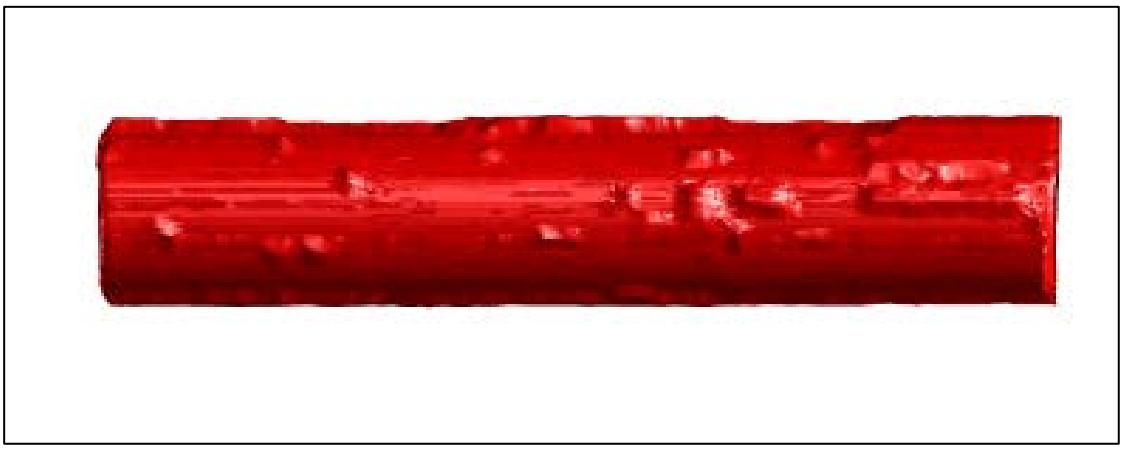


Figure 5-72. 3D reconstruction of pipe surface from multiple frames

A more efficient design of the projector is shown in **Figure 5-73** where all the components are integrated into a compact 3D printed case. In this prototype, all the lenses and slides have been put inside frames that can be moved along a rail so that the components can be inserted and removed easily. This new design offers a better way to easily change the projected pattern without the need to reassemble the whole light module

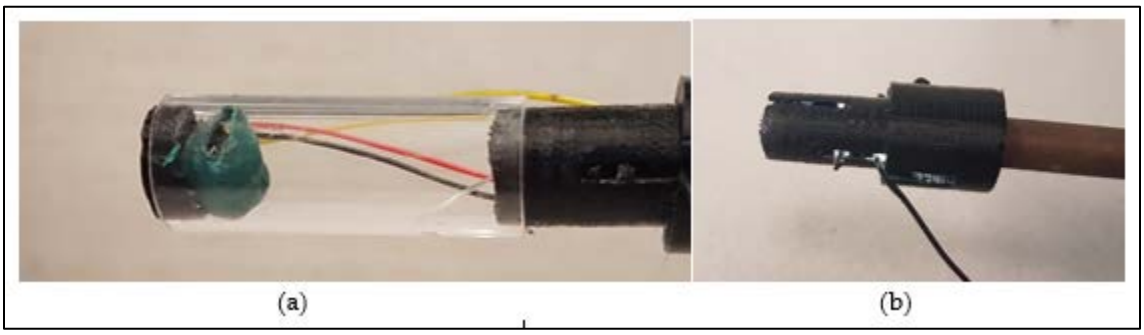


Figure 5-73: The final version of the scanning tool and a separate image of the light module

Damage classification and data reduction

The proposed structured light scanning platform creates high-quality inner wall optical images. These images record the specific condition of the pipeline. How to analyze these images will influence the accuracy and the efficiency of the detection platform. Developing a proper pipe data analysis algorithm is necessary. However, the large amount of the images brings challenges to extract defect features efficiently. Thus, data reduction should be done to reduce data dimensionality and choose significant features firstly. In order to extract pipeline inner wall damage-sensitive features, the histogram of oriented gradients (HOG) descriptor [49] is utilized to reserve image edge information which can represent the deformation features of the structured light. This HOG descriptor is based on calculating well-normalized local histograms of image gradient orientations in a dense grid. The calculation of the HOG descriptor can be summarized as follows:

1: Compute the pixel-level gradient image for each image. Record every gradient's amplitude and orientation. The most common method is to apply the 1-D centered, point discrete derivative mask in one or both of the horizontal and vertical directions. Specifically, this method requires filtering the color or intensity data of the image with the following filter kernels:

$$[-1,0,1] \text{ and } [-1,0,1]^T$$

2: Divide the image into small connected regions called cells, and for each cell compute a histogram of gradient directions or edge orientations for the pixels within the cell. Each pixel within the cell casts a weighted vote for an orientation-based histogram channel based on the values found in the gradient computation. The cells themselves can either be rectangular or radial in shape, and the histogram channels are evenly spread over 0 to 180 degrees or 0 to 360 degrees, depending on whether the gradient is "unsigned" or "signed".

3: Groups of adjacent cells are called blocks. Normalized group of histograms represents the block histogram. In block normalization, L2-norm is used in this report:

$$f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}}$$

v denotes the non-normalized vector containing all histograms in a given block. The set of these block histograms represents the descriptor. **Figure 5-74** demonstrates the HOG descriptor implementation scheme. The 'cells' can characterize defect appearance and shape. The normalized 'blocks' is for better invariance to illumination, shadowing, etc. HOG descriptor reduces the initial data dimensionality by mapping gradients into angular bins in step 2.

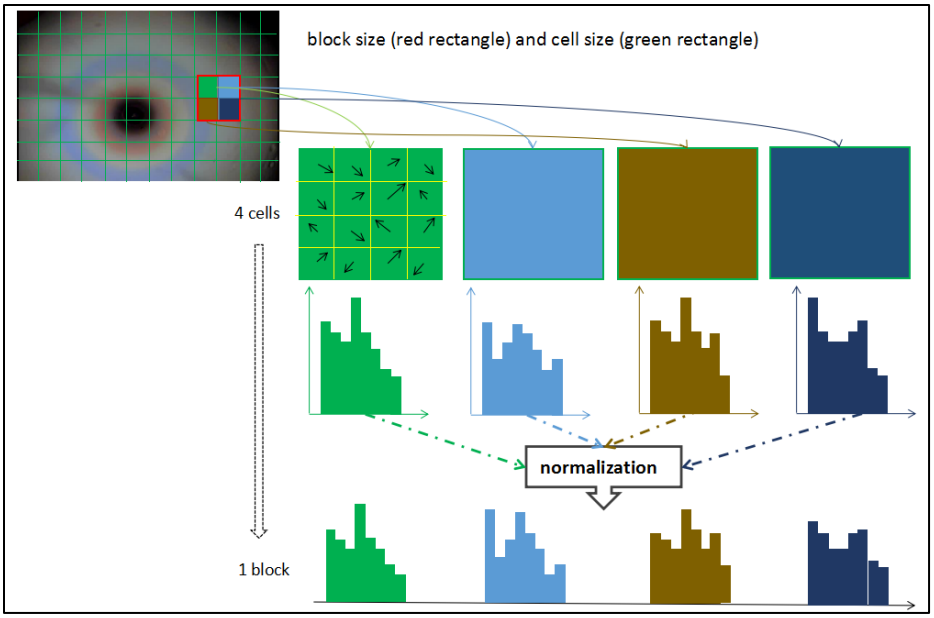


Figure 5-74. HOG descriptor implementation scheme

Moreover, so as to test whether HOG descriptors is appropriate in structured light data feature extraction, linear support vector machines (SVM) algorithm is used to classify the computed HOG descriptors. An SVM model is a representation of the samples as points in space, mapped so that the samples of the separate categories are divided by a clear gap that is as wide as possible. New samples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. Define a training data set of N HOG descriptors $\hat{x}_i$ (i=1,...,N) of the form $(\hat{x}_1, y_1),\ldots,(\hat{x}_N, y_N)$

Where y_i are labels. In this report, y_i=1 means $\hat{x}_i$ contains defects and y_i=-1 means $\hat{x}_i$ contains non-defects. In binary classification problem, linear SVM algorithm is trained by the minimization of the error function:

$$\begin{aligned} &\frac{1}{2}w^T w + T\sum_{i=1}^{N}\upsilon_i, \\ s.t.\ &w^T x_i y_i + b \geq 1 - \upsilon_i, \\ &\upsilon_i \geq 0, \mathrm{i} = 1,\ldots,\mathrm{N} \end{aligned}$$

Equation 5-11

Where υ_i are slack variables which penalizes samples which violate the margin requirements. In order to judge the classification results, the accuracy α is defined as follows:

$$\alpha = \frac{CP + CN}{P + N}$$

Equation 5-12

where CP denotes number of images which include defects correctly classified, CN represents number of images which have no defects correctly classified, P is a number of defect images and N is the number of non-defect images.

Classification results based on HOG descriptor

(1) Simulation pipe data classification

Four types of damages include bump, dent, hole and squeezing defect were simulated in the form of a video (**Figure 5-75**). The size of each frame is 768*1024. Define the size of the cell as 8*8, and then four cells concatenate one block. Nine bins which are spaced over 0◦–180◦ map the pixel-level gradients into nine directions. After these definitions, every image can get their corresponding HOG descriptors. Figure 5-76 illustrates four normalized histograms which separately correspond to four adjacent cells in one block. The HOG descriptor of one image can be evaluated by cascading all blocks. Based on this data set, we can see that the size of each data feature-HOG descriptor is reduced to 36*12065=434340 when the initial image dimensionality is 768*1024=786432. The row number 36 is the product of 9 bins and four cells in one block. 12065 is computed by the step size and the block size because the block scans the whole image. That's to say; the original data has been reduced. Then the SVM classifier is used to test whether the HOG features are meaningful to classify the damages.

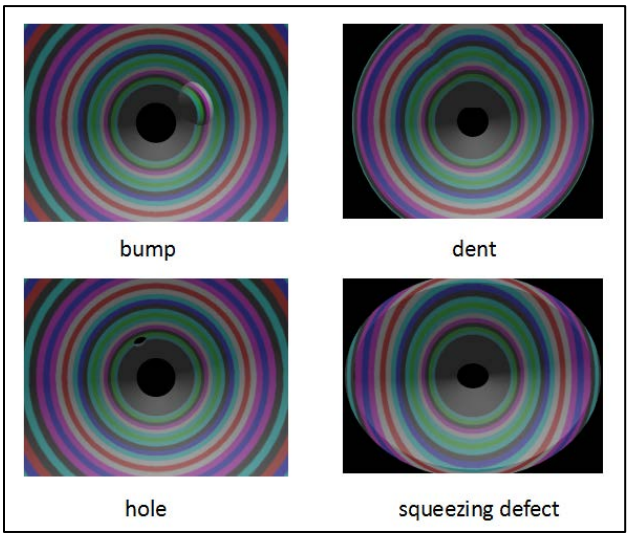


Figure 5-75. Four types of damages

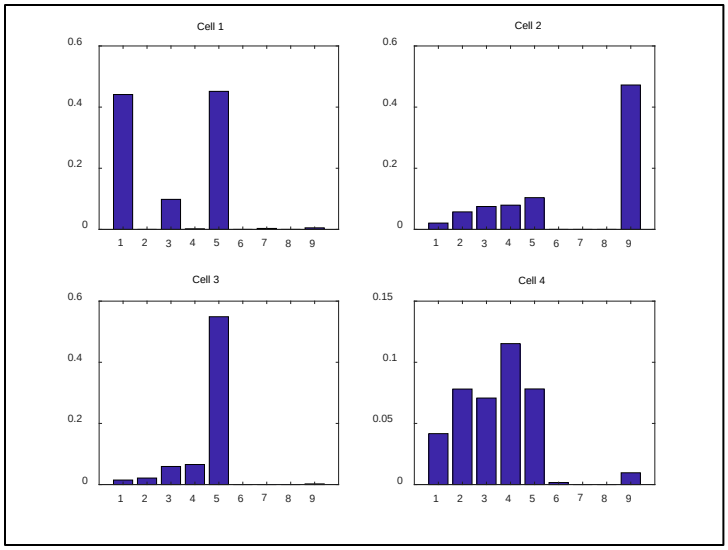


Figure 5-76. Four normalized histograms in one block of one simulation pipe image

Firstly, the SVM classifier is a supervised learning algorithm. One training set with labels should be prepared. Label 1 denotes the defect frame and label -1 means the non-defect sample. In this report, 50 labeled frames with 4 types of damages are used to train the SVM classifier. After training, 26 frames with different types of damages are tested. From the testing results in **Figure 5-77**, 26 samples have been successfully classified. According to the definition of accuracy, the classification results achieve 100% accuracy. That's to say, the HOG descriptors can be used to extract structured light data features.

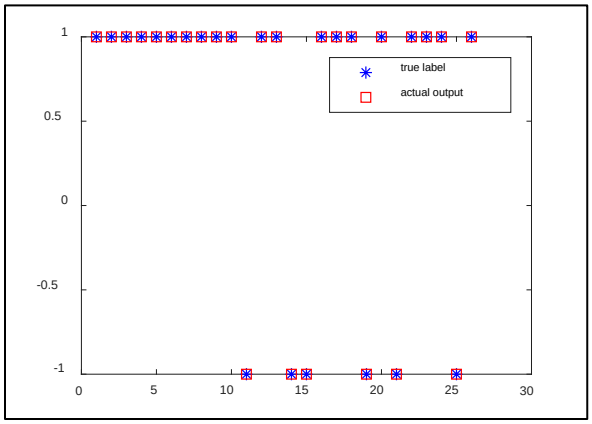


Figure 5-77. The classification results of the test simulation samples

(2) Experimental pipe data classification

The experimental pipe data (480*720) which contains several hole damages are recorded by the Prototype IV scanning system (Multi-Rings multi-Colors scanning sensor). In the same way, all the frames can be represented by their corresponding HOG features. **Figure 5-78** shows four normalized histograms in one block. The size of each HOG descriptor is

36*5251=189036 which is less than 480*720=345600. Separate all frames into one training data set (80 samples) and one test data set (25 samples). From the classification results in **Figure 5-79**, we can see that all test samples have been classified correctly. Combined with the simulation results, it can be summarized that the HOG feature can be utilized to do the structured light data reduction and feature extraction.

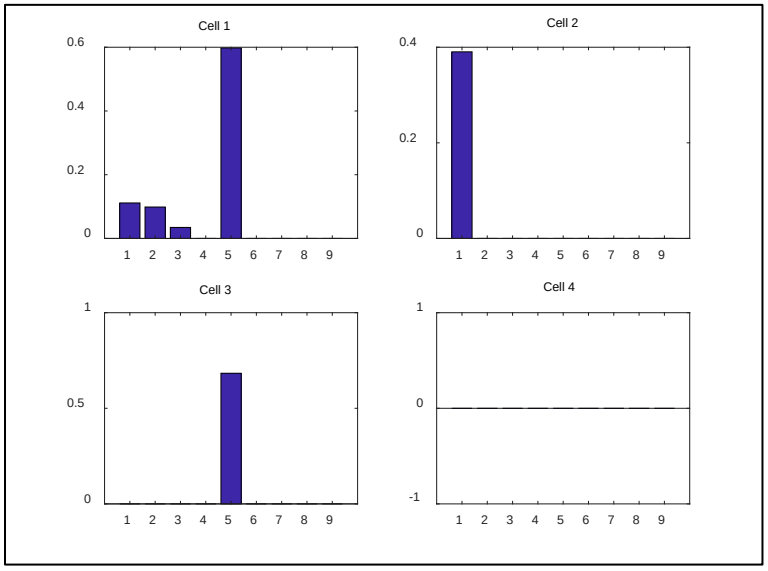


Figure 5-78. four normalized histograms in one block of one experimental pipe image

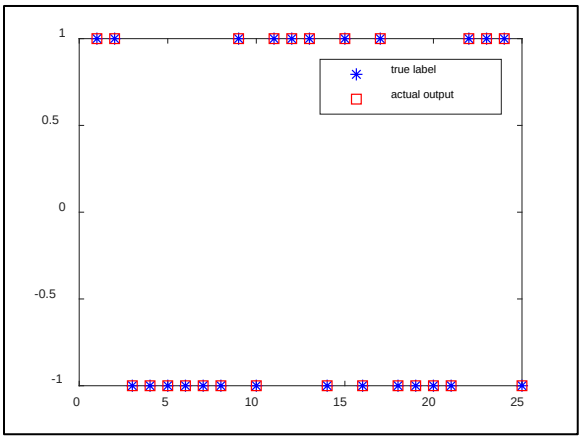


Figure 5-79. classification results of the experimental samples

Pipe data convolutional neural network (CNN) classification model

Data reduction extracts useful features from the initial data which can train the classification model more easily. One deep learning approach - convolutional neural network can learn data features by itself without manual intervention. In order to develop more data processing method to meet the different demands of practical application, another classification model based on CNN is researched.

Background

The inverse problem in NDE about evaluating defect information $\Omega(i)(i=1,2,...,I)$ from observed signals $\Theta(i)(i=1,2,...,I)$ can be denoted as:

$$\Omega(i)=\{M^{-1}\}\Theta(i); \qquad i=1,2,...,I \qquad \textbf{Equation 5-13}$$

where M^{-1} is the inverse operator. The solution of the inverse problem is an estimation $\hat{\Omega}(i)(i=1,2,...,I)$ of defect information $\Omega(i)(i=1,2,...,I)$. One classical CNN structure is developed to solve the evaluation of $\hat{\Omega}(i)(i=1,2,...,I)$.

A whole CNN is comprised of one input layer, one or more convolutional layers, several pooling layers and one output layer. The convolutional layers are feature extraction layers which can compute input's convolutional features without artificial definition and operation. In this layer, weight sharing is applied to increase learning efficiency by reducing the number of free weights being learnt. The pooling layers can reduce the spatial size of the input representation to make the CNN easier to train. Besides these layers, dropout method can be used to prevent CNN from overfitting. In the training phase, the output of the l+1 convolutional layer $\Theta_{m,n}^{labeled,l+1}(i)$ can calculate by the following equation:

$$\Theta_{m,n}^{labeled,l+1}(i)=\sum_{a}\sum_{b}w_{a,b}^{l+1}\sigma(\Theta_{m-a,n-b}^{labeled,l}(i))+b_{m,n}^{l+1} \qquad \textbf{Equation 5-14}$$

where m and n are pixel's coordinate. $w_{a,b}^{l+1}$ is the weight of the l+1 convolutional filter on coordinate (a, b). σ is the activation. $b_{m,n}^{l+1}$ is the bias. Back propagation based on loss function such as mean squared error, categorical cross entropy is a method to calculate the gradient of the loss function with respect to the weights $w_{a,b}^{l+1}$ and bias $b_{m,n}^{l+1}$ in a CNN. The computed gradient of loss function is used to adjust the weights $w_{a,b}^{l+1}$ and the bias.

Pipe data preprocessing procedure

In the designed laser scan and structured light detection system, the deformation of the laser or the structured light which is caused by the defects in the pipe has been recorded in the video. Separate the video into the images and then preprocess them for improving the CNN performance. In order to pay attention to the deformation, one strategy is used to prepare the inputs of the CNN.

One threshold χ is defined to suppress the background (the region that is outside of laser ring or structured light). The inputs of the CNN in the training phase should contain defect features and non-defect features. These features should help training a CNN model which has the invariance like scale invariance and direction invariance. In this prediction model, the local features which is captured by a window Win with fixed size (W_r, W_c) are utilized (red line in Figure 1). The fixed window can guarantee the inputs with the same size and scale. Moreover, because the laser ring or structured light detection system use circular light, image rotation is utilized to make the CNN prediction model be direction invariance (**Figure 5-80**).

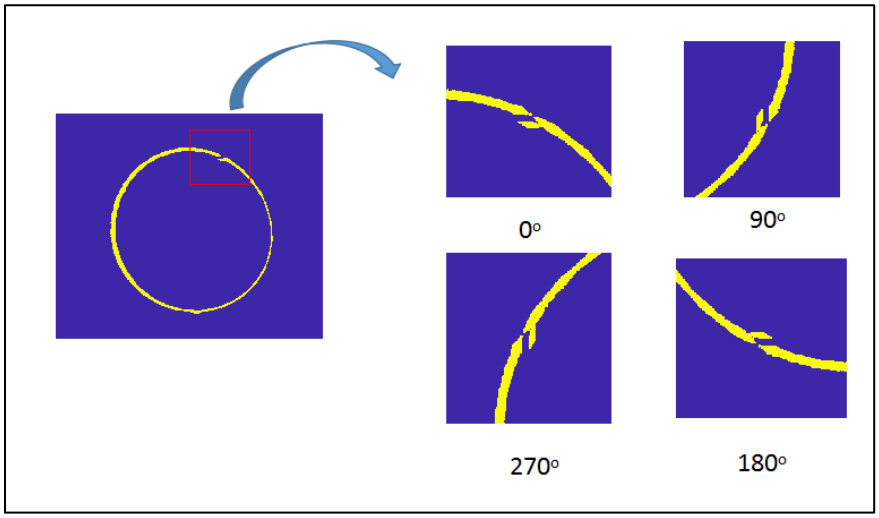


Figure 5-80. image rotation

Pipe data classification results

After data preprocessing, one classical CNN structure which contains four convolutional layers and two pooling layers has been used (**Figure 5-81**). The size of convolutional filters is 3*3. 3 filters utilized in the first convolutional layer and 6 filters in the second convolutional layer. Dropout method is used to prevent the overfit of the CNN model. The output layer is a softmax layer which can map the outputs of samples into the probabilities.

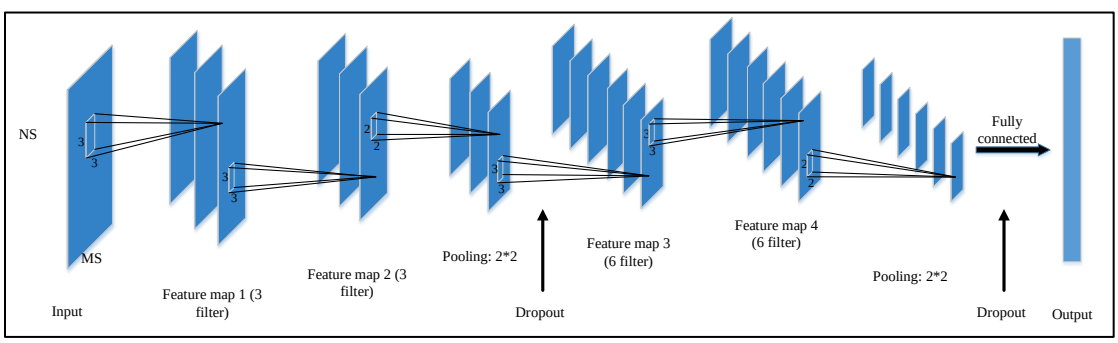


Figure 5-81: CNN structure

Specifically, in this qualitative classification problem (defect or non-defect), defect information $\Omega(i)(i=1,2,\ldots,I)$ can be regarded as an indicator set like [0, 1] that indicates whether defects exist or not. In general, the outputs of the designed model are their inputs' corresponding labels:

$$\Omega(i)=\begin{cases} 1, \textit{defect} \\ 0, \textit{non-defect} \end{cases}, \qquad i=1,2,\ldots,I \qquad \textbf{Equation 5-15}$$

After the SoftMax layer, these labels will be mapped into values between 0 and 1. If the output is close to 1, it means this sample contains defect. One threshold τ is needed to determine which class they belong to:

$$\Omega(i)=\begin{cases} \geq \tau, \textit{defect} \\ < \tau, \textit{non-defect} \end{cases}, \qquad i=1,2,\ldots,I \qquad \textbf{Equation 5-16}$$

The laser ring data is used to test the performance of the CNN model. Thirteen hole defects on the plastic pipe is used to train the CNN model. The laser scans the plastic pipe continuously. So, one defect information will be recorded in a consecutive sequence. **Figure 5-82** shows an example of one testing frame. Define W_r=200, W_c=200. The fixed window Win segments this image (after background suppressing: χ =240) into 16 parts. Record the rank of every part (as shown in Figure 5-82), the region of defect can be known approximately after the testing. Figure 5-83 displays the classification results of the testing samples. A consecutive sequence of 7 frames have been put into the trained CNN model to test the performance of the CNN model. This sequence has recorded the same defect. As shown in **Figure 5-83**, only part 3 in every image contains defects. The output of the CNN model is a softmax function, so the output of each sample is a probability. Combined with the labels we defined as below, if the output of one sample is close to 1, it indicates that this is a defect sample, vice versa. Compared this information with the probability results in **Figure 5-83** , it is obvious that all defect areas have been identified. Define $\tau=0.5$. All part 3 of 7 frames have larger probabilities than 0.5. Other parts are smaller than 0.5 which means they have non-defects. All defect regions have been successfully identified without false calls.

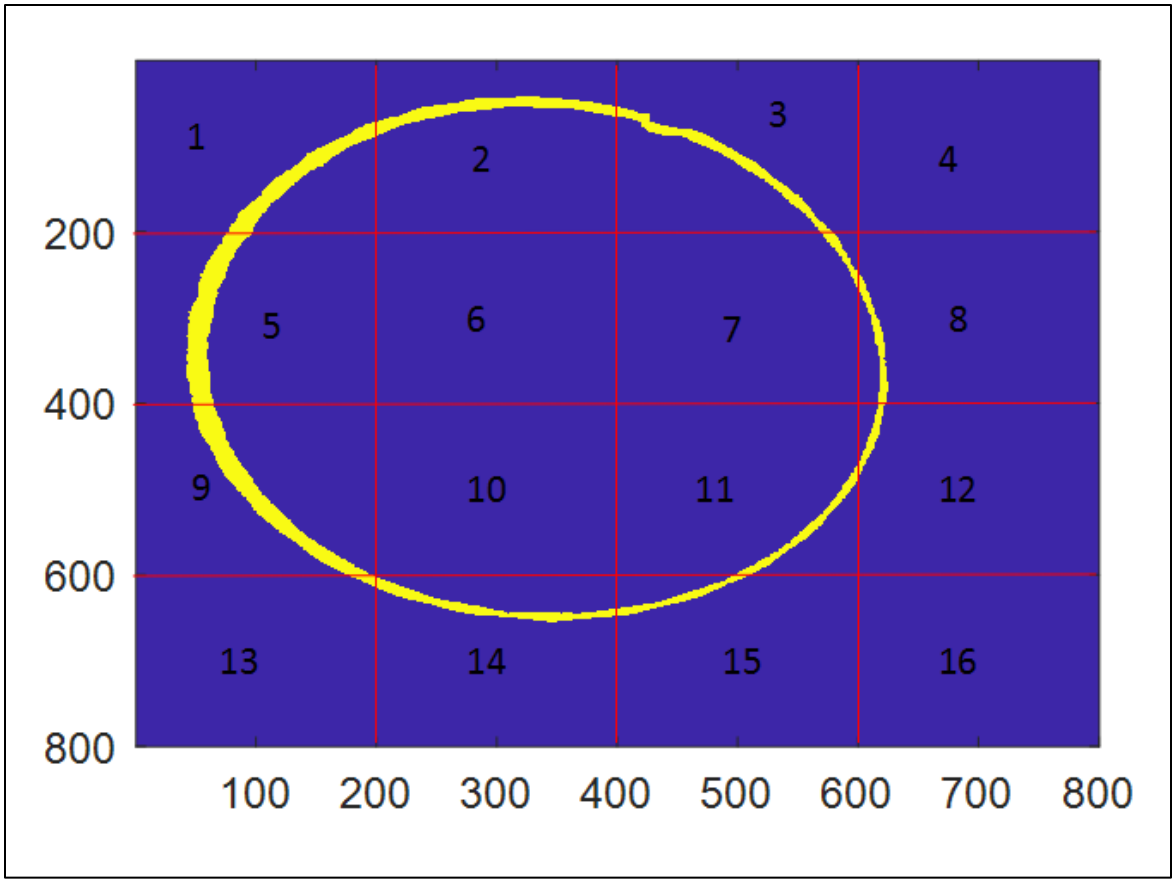
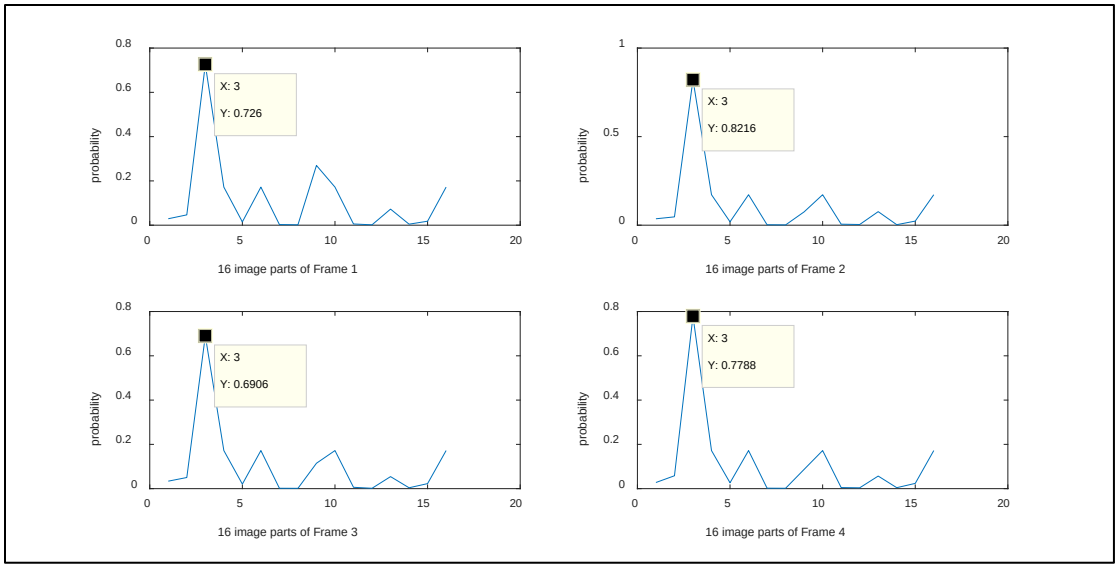
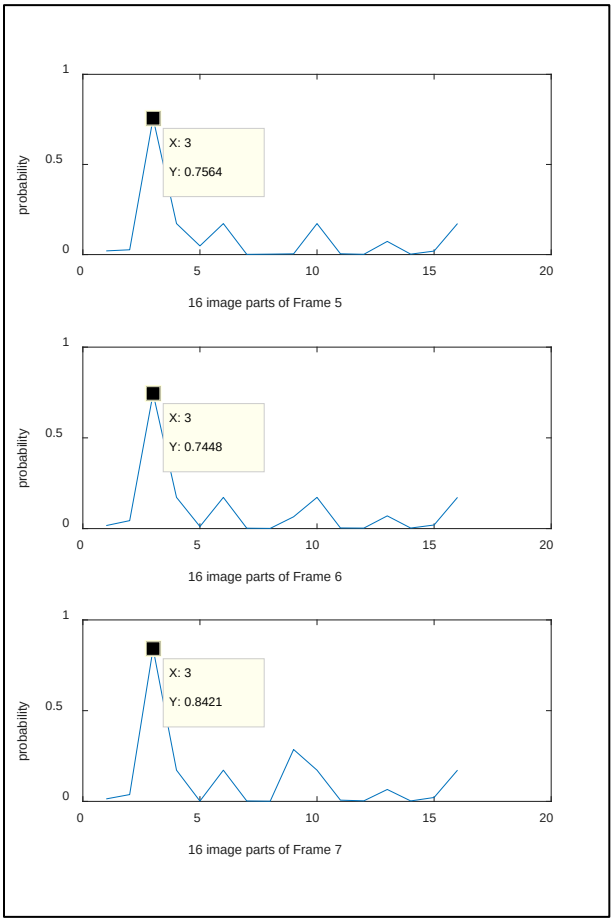


Figure 5-82. segment result of one image



(a)



(b)

Figure 5-83. classification results of a consecutive sequence of frames

Conclusions

The project aims to develop an integrated set of quantitative tools that provides an approach to reducing operational risk in vintage plastic distribution systems susceptible to Slow Crack Growth (SCG) failures. The tool is an optical device that can reconstruct the profiles of the internal pipe surface and able to detect the existence of deformations and defects that can be visually detected on the pipe wall. For this task, structures light endoscopes are proposed due their robustness and simplicity. In this project, the experimental work is divided into two parts. The first part is used to develop and test the performance of laser profiling based scanners while the second part is specified to develop and test a multicolor multi ring structured light scanners. Four prototypes have been successfully developed for the high resolution and sensitivity pipe inner-wall imaging. The first design is developed with an initial capability to give a partial image about the pipe surface. This initial prototype is used to help developing the reconstruction algorithm and became the base for the next generations of the prototypes. The reconstruction algorithm is developed and tested on simulation data and then used for the reconstruction of the images from the actual laser scanner. In order to enhance the quality of the device and increase the amount of the reconstructed surface area, prototype II is developed. This prototype is 3D printed and has employed a new fishery camera to provide the required wide angle to capture the full laser ring. This prototype provides a more stable scanning platform that can reconstruct 100% of the pipe surface. In order to further improve the quality of the scanner, multicolor multi-ring approach is used to increase the number of reconstructed points from each frame and enhance the spatial resolution by providing sharper edges. In order to project colored rings into the pipe a small slide projector, prototype III is developed. The projector is built to project two-color rings. Prototype IV is developed to reduce the size of the scanner in order to be fitted inside a 1-inch pipe and to increase the number of projected rings. The prototype has been built and tested, and a simulation environment is created to develop the reconstruction algorithm. Experimental scans have been performed and the data is used to create the accurate profile of the pipe inner wall. The performance of the algorithm has been tested against different defect types by using simulations. Smooth deformations and shape changes were reconstructed correctly and precisely. Sudden discontinuities in the surface introduce non reconstructable shadows behind the defect and also create difficulties in the matching between the projected and received sequences. The code discontinuities are not a severe factor because as the scanning platform is moving along the object, it can collect all the missing points. The shadow problem is dependent on the view angle of the scanner (larger view angle creates less shadows) and the smoothness of the deformation. A synthetic database is created to act as a reference for our data analysis algorithms. The database includes five defects types with different shapes and sizes and it is mainly used for the data reduction analysis. A data reduction framework is created and followed by a classification network. The network model works as a binary classifier that test each set of camera frames for the presence or absence of defects. From the classification results, it can be seen that the

classification model performs excellent in the defect prediction problem and the defined local feature can reserve the characteristics of the initial data. The data reduction model is also helpful to improve the performance of the classification network.

Future work

The delivered work meets the goal and proposed research objectives. Some additional research and hardware development can be done to enhance the ESLiST sensing system.

1. The current system provides a high quality and fast surface profile reconstruction, but the spatial resolution can be further increased by improving the hardware design, better system calibration, and more efficient reconstruction algorithm. Currently, a fixed structured light pattern was applied, which might not be optimal for all damage types and different field testing conditions. With the successful demonstration of the Prototype, I to Prototype IV and their feasibility of extracting damage dimensional information at MSU, extensive involvement of GTI in collecting real data should be expected. Additional work on feeding updated stress concentration factor derived from ESLiST image and ESLiST assisted decision information back to the sensing system should be carried out to complete the development loop and further optimize the novel structured light sensors design.
2. The proposed prototype works very well in a controlled lab environment by assuming the sensor is moving in a controlled straight path that is along the pipe's center axis. This assumption might be violated in actual field testing and inspection due to mechanical vibration of the sensing platform moving inside the pipe with complex geometries. An inertial measurement unit (IMU) can be integrated into the system to provide an estimation of the device orientation and positioning. Real-time compensation algorithms are needed to correct any distortion and misalignment induced error in data reconstruction to achieve same imaging resolution in a field environment.
3. Static structured light patterns have been studied and implemented as we proposed. However, the potential of the structured light approach will be maximized by introducing dynamic light patterns and data structures to adapt to various damage types and sizes under different light conditions. This could be better understood through both numerical simulations and experimental studies by collaborating with GTI and industry partners. An optimization of the current static structure light patterns to study the relationship between the smallest detectable damage feature and the parameters of the static light pattern will also be beneficial.
4. Further improvement of the current data reduction and reconstruction will lead to faster and more accurate damage detection. The current framework only focuses on the SCG failure mode, so supervised multispectral data dimensionality reduction and defect classification methods would be sufficient and are successfully demonstrated. However, additional work should be done to realize the unsupervised damage recognition and information fusion while multiple failure modes are considered. Integration of advanced sensors from MSU and the intelligent decision support system from GTI and ASU would be more critical to achieving an optimal diagnostic and prognostic framework for gas pipeline industries.

Section 5 Bibliography

27. Boenisch, A. *Advanced subsea pipeline inspection: an overview of flied experience and benefits of subsea pipeline*. in *Pipelines International*. 2013.
28. Barnwell, C., W. Doyle, and C. Gunn, *LiDAR for terrain mapping on the Alaska pipeline corridor.* Office of Federal Coordinator, accessed at http://www. arlis. org/docs/vol1/AlaskaGas/Paper, 2009.
29. Duran, O., K. Althoefer, and L.D. Seneviratne, *Automated pipe defect detection and categorization using camera/laser-based profiler and artificial neural network.* IEEE Transactions on Automation Science and Engineering, 2007. **4**(1): p. 118-126.
30. Zhang, L., B. Curless, and S.M. Seitz. *Rapid shape acquisition using color structured light and multi-pass dynamic programming*. in *3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on*. 2002. IEEE.
31. Drouin, M.-A. and J.-A. Beraldin, *Active 3D Imaging Systems*, in *3D Imaging, Analysis and Applications*. 2012, Springer. p. 95-138.
32. *CCTV Pipeline Testing Survey & Reporting with Wincan v8*. 2013]; Available from: http://www.auscamqld.com.au/.
33. Agency, I.A.E., *Training Guidelines In Non-Destructive Testing Techniques*. 2013, International Atomic Energy Agency.
34. Newsletter, S.C.G. *Applications of Computer Vision to Computer Graphics*. 1999 November 1999]; Available from: https://www.siggraph.org/publications/newsletter/v33n4/contributions/curless.html.
35. Gonzalez, R.C., S. Eddins, and R.E. Woods, *Digital image processing using MATLAB*. 2004.
36. Salvi, J., et al., *A state of the art in structured light patterns for surface profilometry.* Pattern recognition, 2010. **43**(8): p. 2666-2680.
37. Geng, J., *Structured-light 3D surface imaging: a tutorial.* Advances in Optics and Photonics, 2011. **3**(2): p. 128-160.
38. Monks, T. and J.N. Carter. *Improved stripe matching for colour encoded structured light*. in *International Conference on Computer Analysis of Images and Patterns*. 1993. Springer.
39. Boyer, K.L. and A.C. Kak, *Color-encoded structured light for rapid active ranging.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987(1): p. 14-28.
40. Guan, C., L. Hassebrook, and D. Lau, *Composite structured light pattern for three-dimensional video.* Optics Express, 2003. **11**(5): p. 406-417.
41. Schmalz, C. and E. Angelopoulou. *A graph-based approach for robust single-shot structured light*. in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. 2010. IEEE.
42. Schmalz, C., et al., *An endoscopic 3D scanner based on structured light.* Medical Image Analysis, 2012. **16**(5): p. 1063-1072.
43. Commons, W. *Ellipse Properties of Directrix.svg*. Available from: https://commons.wikimedia.org/wiki/File:Ellipse_Properties_of_Directrix.svg.
44. Pi, R. *Raspberry Pi 2 Model B*. Available from: https://www.raspberrypi.org/products/raspberry-pi-2-model-b/.

45. Young, M., et al. *coded structured Light*. in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. 2007. IEEE.
46. Turk, G. and M. Levoy, *The stanford bunny*. 2005, Stanford University Computer Graphics Laboratory, URL http://graphics.stanford.edu/data/3Dscanrep.
47. Je, C., S.W. Lee, and R.-H. Park. *High-contrast color-stripe pattern for rapid structured-light range imaging*. in *European Conference on Computer Vision*. 2004. Springer.
48. Zhang, Z., *A flexible new technique for camera calibration.* IEEE Transactions on pattern analysis and machine intelligence, 2000. **22**(11): p. 1330-1334.
49. Lin, Z., M. Chen, and Y. Ma, *The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices.* arXiv preprint arXiv:1009.5055, 2010.

6. Damage Detection from Optical Imaging, Creep-Crack Growth Prediction, and Condition-based Maintenance Framework for Aldyl-A Pipes

This section will summarize all work done by ASU during the project. A novel damage detection method based on optical imaging were developed. Damage classification was done using a Naïve Bayes (NB) classifier. The classification result was compared with the neural network (NN) and showed little difference. In searching for an algorithm to build a new classifier that could encode empirical information (expert opinion) into the classifier, a novel Bayesian/maximum entropy network (BEN) classifier were developed based on the maximum entropy concept. The BEN classifier can handle extra information such as moment constraints. Given the available information, the classifier can achieve faster learning and behaves better than NB when the training size is small.

An equivalent creep crack growth (CCG) prediction model using a power law equation was calibrated and validated using the experimental data from GTI. With the proper assumptions, the model prediction matches the experimental data reasonably well. Uncertainty was introduced into the model parameters and Monte Carlo (MC) simulation was used to calculate the failure probability. The model will be later applied into the maintenance framework.

A condition-based maintenance framework was formulated for the maintenance scheduling of the pipeline system. In the proposed framework, the conditions of the pipes were categorized into several condition stages according to their crack length. The conditions of the pipes can be iteratively calculated via a condition transition matrix. The condition transition matrix can be calculated using the predicted crack growth curve. The overall maintenance plan is optimized by minimizing the maintenance cost under the constraint of failure probability. The maintenance was done using genetic algorithm.

The model parameter in the CCG prediction model can be updated by observations via Bayesian updating. The updating process would reduce the uncertainty within the model parameters and hence reduce the variances. The updated CCG curve would change the values in the probability transition matrix in the maintenance framework and hence change the maintenance decision. It is shown that, comparing with non-updating results, the cost for maintenance was reduced.

Introduction

Damage diagnosis and remaining life prediction of pipeline infrastructure systems is still a challenging problem despite tremendous progress made during the past several decades in understanding damage accumulation in plastic gas distribution pipes. Historically, two entirely different approaches are used for structural system performance prediction (i.e. data-driven or physics-based predictive models). Data-driven approaches used nondestructive inspection technique (optical images, ultrasound, acoustic measurement, etc.) and experts' justification (personal experience on trending function, normal range of operations, etc.) to extrapolate system future behaviors. Physics-based models use underlying mechanisms (crack initiation and propagation model, chemical diffusion functions, oxidization rate, etc.) to predict system future behaviors. Information fusion between two approaches will enable accurate risk assessment and mitigation planning.

Bayesian networks are graphic probabilistic models that are based on Bayes' theorem [50]. Bayesian networks describe the dependence between variables through a direct acyclic graph (DAG). Bayes' theorem states that the posterior probability is proportional to the product of the likelihood and the prior [51]. Bayesian networks are widely used in inference due to the ability to update the posterior distribution with observations. A lot of work has been done on pattern recognition [52, 53], image classification [54] and automatic image segmentation [55]. There are different kinds of algorithms for Bayesian network classifiers, such as Naïve Bayes, Tree Augmented Naïve Bayes (TAN) [56] and Selective Naïve Bayesian Networks [57]. The Naïve Bayes classifier is a simple and fast algorithm with a surprisingly good accuracy despite the unrealistic assumption about the independence of the features [56]. In this project, a device to reconstruct pipe inner surfaces in 3D using video frames was built. The device comprised an endoscope camera and a patterned projector. Video was recoded as the device move along the pipe. Reconstruction was done by the analysis of the video frames and triangulation of the patterned light in each frame. The idea was presented in some medical researches [42]. The device building and reconstruction algorithm was mainly done by the MSU. Based on the reconstruction result, the geometric features such as the length, area of the damage can be isolated and calculated. These features were put into a Bayesian classifier. The classification result was compared with more advanced algorithms like the Neural Network (NN) and showed no big difference in our case.

To further enhance the accuracy of the algorithm based on the Naïve Bayes (NB) classifier, various methods have been developed. The selective Naïve Bayes classifier method uses a forward search to select features that does not decrease the accuracy. An approach of combining the Naïve Bayes with decision tree method was proposed in [58] and showed significant improvements. TAN adds additional edges between features to capture their correlations and releases the strong independence assumption of Nave , but does not significantly increase the computational cost. Although tremendous progress has been

achieved in machine learning and artificial intelligent over the decades, a classifier would still need a large dataset to be trained and would only take point data (instances). It is hard however, to encode empirical information (constraints over a feature) into the classifier. In this research, we presented that by introducing an additional term to the Bayesian equation, we can handle extra information in the form of constraint in the proposed Bayesian/maximum entropy network (BEN). The added term was derived using the maximum entropy (ME) method [59, 60] by maximizing the entropy of the posterior distribution under constraints. It was shown that Bayes' theorem is a special case of ME [61]. A demonstration example showed that the BEN classifier behaves better than traditional NB when the training size is small.

Since the working condition of the pipe is under static pressure, creep failure is the dominant failure mode in the pipeline system. Many existing models have been proposed and developed to analyze the creep deformation and life prediction. The characteristic of a creep crack is that in polymeric materials is that there is craze formed at the crack tip, which is fibrils that are bridging the edges of the crack. Schapery's [62] model uses viscoelastic fracture mechanics approach to determine the crack growth rate as a function of stress intensity factor (SIF). Williams and Marshall's model [63] considers the Young's modulus as a time dependent function. In [64, 65], cohesive zone model is used to model the creep cracking process as a one-dimensional zone at the crack tip, which grows and ruptures as the time increases. The above studies indicate that SIF is the key for the creep failure prediction. This work used a power law equation to describe the crack growth rate. The model assumed a semi-circular surface crack in the longitudinal direction and the SIF solution were adopted from [65] and an asymptotic solution [66] to consider the notch effect. Model parameter uncertainty were considered to predict life in a probabilistic sense. The model was calibrated and validated by the experimental data from GTI. The probability transition matrix was calculated using Monte Carlo (MC) simulation.

In the search for state-of-the-art method of developing a maintenance frame work, the most discussed method is a condition-based maintenance(CBM). This method is constituted by 3 parts, data acquisition, data handling and decision making. Data acquisition is a step to collect useful data from the object, including event data and condition monitoring data. The event data include information of jobs or operations the object has been through, like loading, temperature change etc. Condition monitoring data are physical measurements of the object that can identify its health state. Both types of data are equally important [67]. There are numerous ways of data acquisition for the condition monitoring data, like microwave or ultrasonic sensor etc. In our case, the condition data would be the image reconstruction. Data handling is how the acquired raw data can be transformed into useful understandable and easy-to-process data. In our case, the acquired reconstruction data was classified into types of damages. The last step is the maintenance decision making, basically diagnostics and prognostics. Prognostics predict faults before it occurs while diagnostics is important when the

prognostics failed and a fault occurred. There exist large quantities of literature regarding the subject, some researchers use hypothesis testing [68] and cluster analysis [69] for fault diagnosis and some uses machine learning techniques [70] but due to the lack of training data, the real application of machine learning technique is not easy. Other researchers use model based approaches with explicit physical and mathematical model, some models for bearings [71] and gearboxes [72] are already applied in diagnosis. The prognostics mostly predicts the remaining useful life (RUL) of a subject under the current condition and past and future jobs or operations. Some methods predict failure if the condition reaches a threshold value and others uses a model based on failure mechanism. Two statistical models called proportional-hazards model [73] and proportional intensity model [74] has become a useful tool in remaining life predictions. A hidden Markov model can calculate the transition probability from known experimental data [75]. Some also tried to apply artificial intelligence to RUL in [76]. Model-based approaches are presented in [77]. For the maintenance decision making, most researches focus on minimizing the costs, although there are some other papers that discuss the optimization of inspection intervals [78]. This project presents a maintenance framework based on probabilistic models. From the pipe imaging section, we can have the types of damages and its corresponding size (can be regarded as a deterioration stage). Depending on its future operation condition, a transition probability matrix can be formed and used to iteratively calculate the probability of the condition of the pipes. A maintenance decision can be made for minimizing the cost while constraining the failure probability of the system. Or if the budget is fixed, a maintenance schedule can be made for maximizing performance.

Bayesian network updating has been extensively used for damage diagnosis and prognosis of metallic and composite materials [79, 80]. The information fusion between diagnostics and prognostics can achieve a more accurate risk assessment and maintenance planning. To achieve a dynamic maintenance network, the model parameters will be continuously updated through the on-field observation. The updating process is achieved by Bayesian updating. This algorithm has been proved in many studies to improve the prognosis using the observation data from the condition monitoring [77][33]. The Bayesian updating is based on Bayes' theorem to update the hypothesis probability distribution (prior) with new information (likelihood). With the continuous updates, the variance of the distribution would reduce, indicating the uncertainty of the parameter would decrease.

Imaging reconstruction and damage classification

This part will discuss the performed work and results related to the pipe 3D surface reconstruction and the damage classifications. First, the algorithm of a Bayesian classifier will be introduced. Before the prototype reconstruction device was available, simulated pipe imaging data were used to test the classifier and showed promising result. Despite the difficulty using the prototype camera to get good pipe imaging, we tested a small set of real data recorded in lab with the classifier and showed good result.

Naïve Bayes classifier

The Bayesian network classifier is based on a simple Bayes theorem:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \qquad \textbf{Equation 6-1}$$

where A and B are events, and P(A), P(B) are the probability of event A and B without regard to each other. P(A|B) is a conditional probability describing the probability of event A given that B is true, also called posterior. P(B|A) is the likelihood term describing the probability of observing B given that A is true. During the classification, a set of training data is required to train the Bayesian network. For each class, the network treats each feature as random variables and find a certain distribution from the data. For a Naïve Bayes classifier, he classification is done by:

$$c^* = \arg\max_{i=1\ldots m} P(c_j)\prod_{i=1}^{n} P(A_i = a_i \mid c_j) \qquad \textbf{Equation 6-2}$$

It means that given a new data, for each feature A_i, the classifier will calculate the probability that it belongs to class j (j=1…m). And assign the class that has the highest posterior probability for this data. The posterior probability is calculated by the product of the probability for each feature since the Naïve Bayes assume independence among all features. A graphical illustration of a Naïve Bayes network is shown in **Figure 6-1**

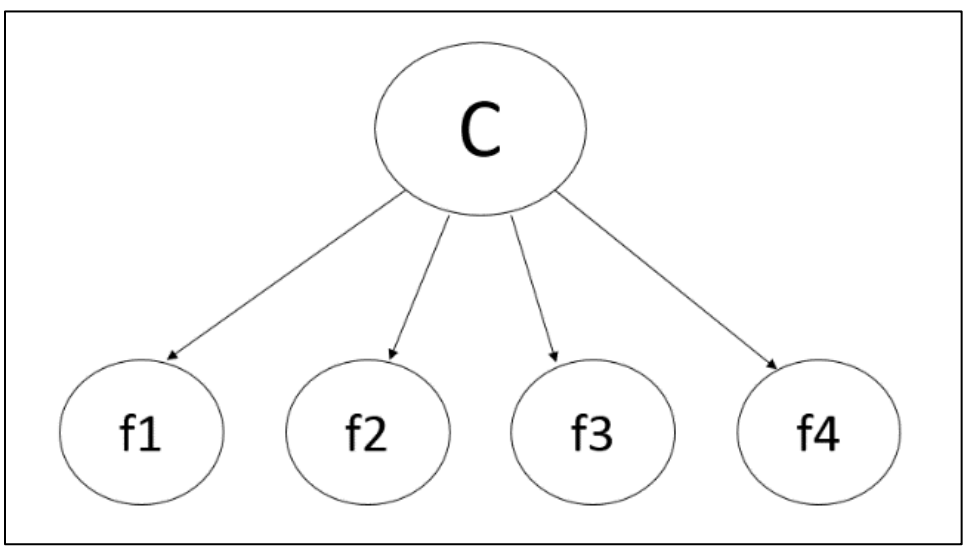


Figure 6-1. Graphic model for a Naïve Bayes classifier with one class node and four feature nodes

The Bayes classifier was widely used since its simplicity in structure and computation.

Pipe imaging data simulation and classification

Pipe simulation

In the pipeline system, common damages include dent, slit, rock impingement and squeeze-off **Figure 6-2**. These 4 types of damage were simulated in a 3D binary array with dimension 640*480*200, where 0 represents the void and 1 is the pipe wall. The damages were generated frame by frame with gradual change in the shape of the circle. The dent and impingement was simulated as an indent on the circle, while dent has a larger deformation and impingement being shaper. The squeeze-off was simulated as a gradual change in the diameter of the cross section of the pipe.

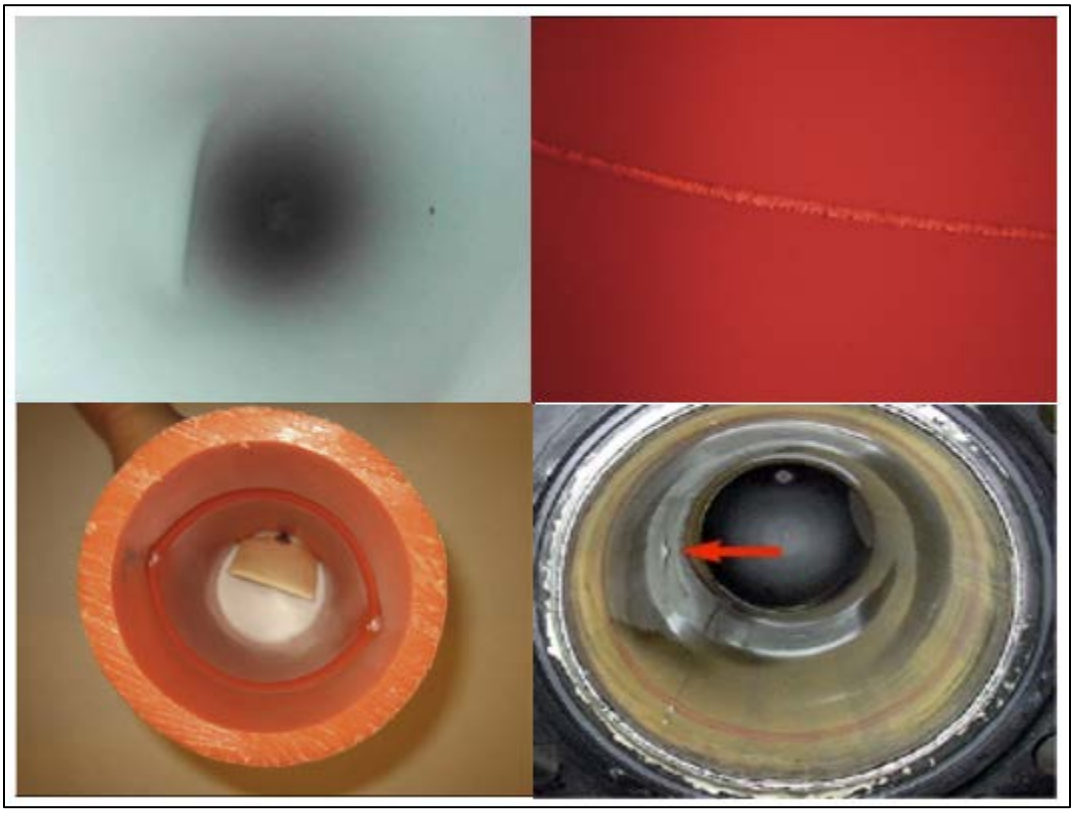


Figure 6-2. Four types of common damage in pipe: dent (top left), slit (top right), squeeze-off (bottom left) and impingement (bottom right)

The four types of damage can be randomly generated with different size and position. This can provide us with great amount of data since the real data is limited. **Figure 6-3** shows a visualization of the 3D array for samples of each damage type. Comparing with the ones in **Figure 6-2**, the simulation can have a good representation of the real damage.

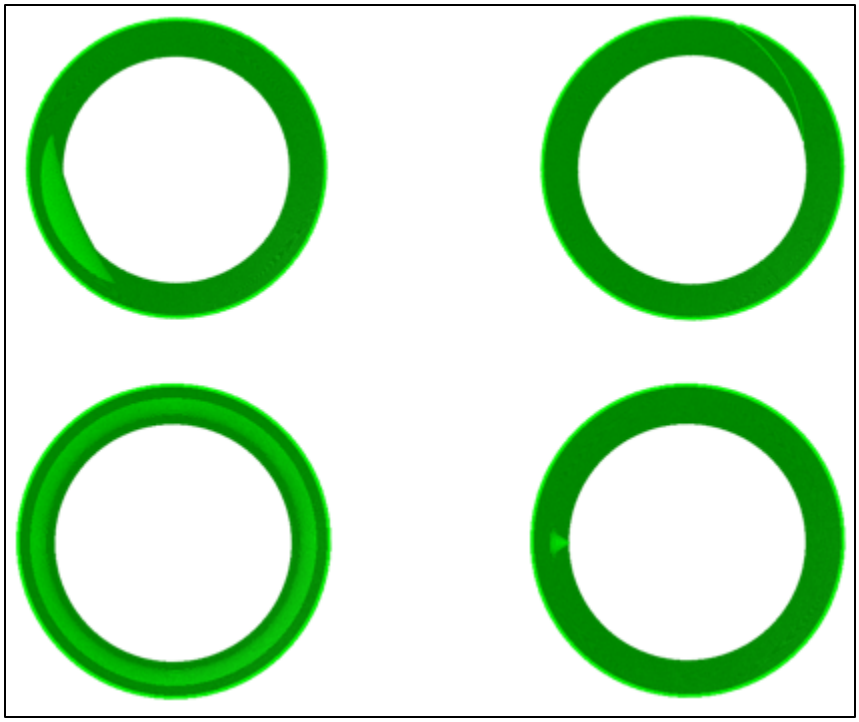


Figure 6-3. . Sample of the four types of damage: dent (top left), slit (top right), squeeze-off (bottom left) and impingement (bottom right)

Random noises were added to the 3D array to simulate the potential noise in a camera sensor.

Feature calculation and classification

Geometric features were extracted based on the simulated pipe. A few methods were tried to isolate the damage in a 3D pipe model. The first method used was a frame differencing method, it regards the previous frame as the background of the current frame. Such method works fine only with idealized data, i.e. the one without noise. A frame averaging method [81] take the average of a few numbers of the previous frames as the background of the current frame. This method would compensate a small amount of noise. But since the camera needs to fit in small pipelines, the size of the sensor would be small and a large amount of noise is expected. A more advanced method of foreground detection using Gaussian mixture model [82] were tried. After comparison, the frame averaging algorithm is chosen as the denoise method in our case due to its low computational cost as well as its good effects on the data.

Based on the isolated 3D structure (damage), a few geometric features can be calculated. In the demonstration example, four features were proposed, namely the surface area of the isolated damage, the maximum cross section (x-y plane) area, the length in z direction and the ratio of x-y plane projection to z direction length, respectively. 400 simulated damaged

pipe sections (100 of each kind of damage) were randomly generated for the feature data. Using these feature, a NB classifier can achieve an accuracy of 89%. Figure 6-4 shows the accuracy of the classifier vs. the training size. The drop of the accuracy at 300 training data size is caused by overfitting.

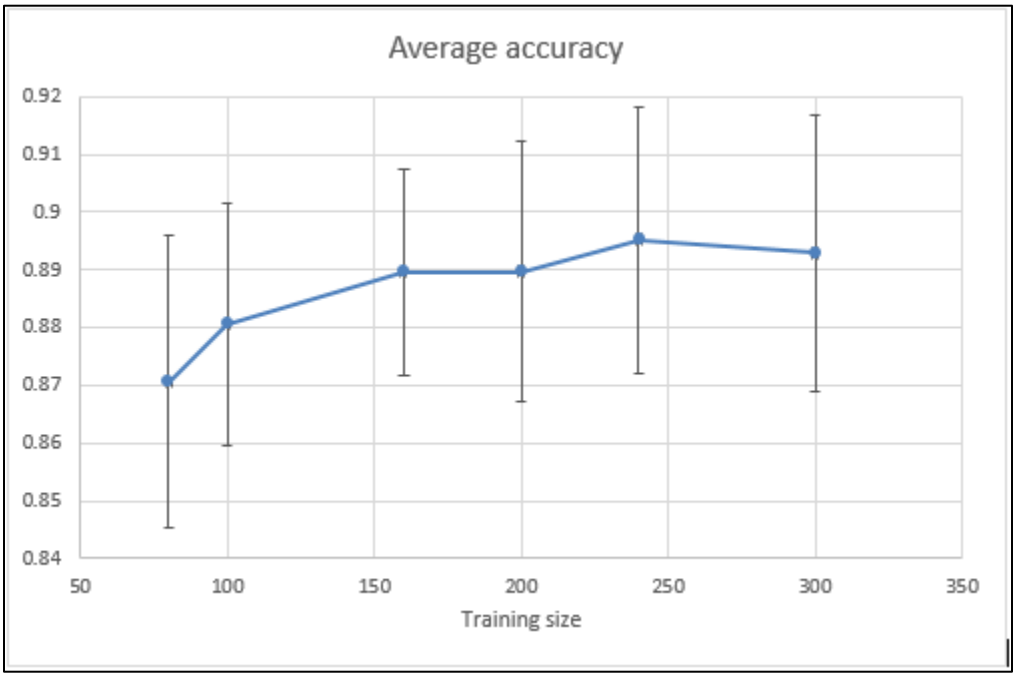


Figure 6-4. Average accuracy vs. training data size with simulated feature data

Comparison with neural network (NN)

The previous section discussed the behavior of the simple naïve Bayes classifier. It can be considered that it achieved plausible results. In this section, a comparison was done using the same data set by classification using Neural Network (NN) in MATLAB [83]. Since NN classifier can directly take image as input, another method of using 2D image as input was introduced and compared among Neural Network and NB.

Neural Network with extracted feature data

Since NN has an additional validation step, the training and testing is a bit different then the NB. The network used in this case has 10 hidden neurons, the training set is set to be 30% to 80% of the data, and the test and evaluation set are set to be equal. The detailed algorithm of NN is beyond the scope of this study. The comparison of the classification result is shown in Figure 6-5. As we can see, the Neural Network has a similar accuracy as NB. And NN also suffers the overfitting effect.

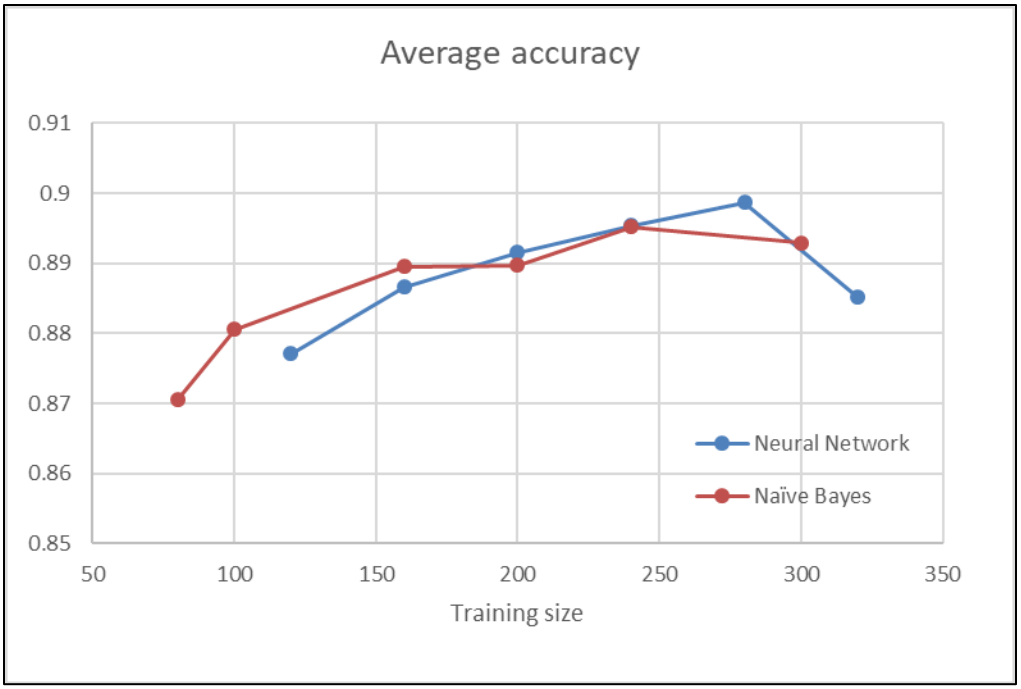


Figure 6-5. Average accuracy for Neural Network using simulated data

2D feature image as input

Although to the benefit that NN can take image as input, the original frames are too large, which would take longer training and testing time, and have too much useless data (both the inside and outside the pipe cross section have black area that is the same no matter the damage). A method of dimension reduction is considered. And a sequence of five steps of image processing is designed to get a feature image with a much lower resolution, which will then act as input images for NN and NB. The targeting image is each frame for the simulated pipe structure with noise. The five steps are dilation, filling, subtraction, rotation and pooling, as shown in **Figure 6-6**. Details will be introduced bellow.

The first step is to dilate the white pixels. This will let the noisy edge of the pipe form a closed contour of the cross section. In the second step the closed section is filled with white pixels, as shown in **Figure 6-6** (c). Then the filled cross section is subtracted from a perfect cross section of the pipe. This step isolates the damage. Since the damages from each frame is at a different angle, a rotation helps align them in the same direction. **Figure 6-6(**e) shows an example of 50 dent section that were aligned in the same direction. Then a pooling zooms in at the damage, forming a 60 by 60 feature image for each frame, further reduce useless information.

Since this method uses only 2D image frames instead of 3D structure, it is able to get much more data set from the simulated pipes. In total, one thousand images were processed with 200 for each damage and 200 for non-damaged pipe section.

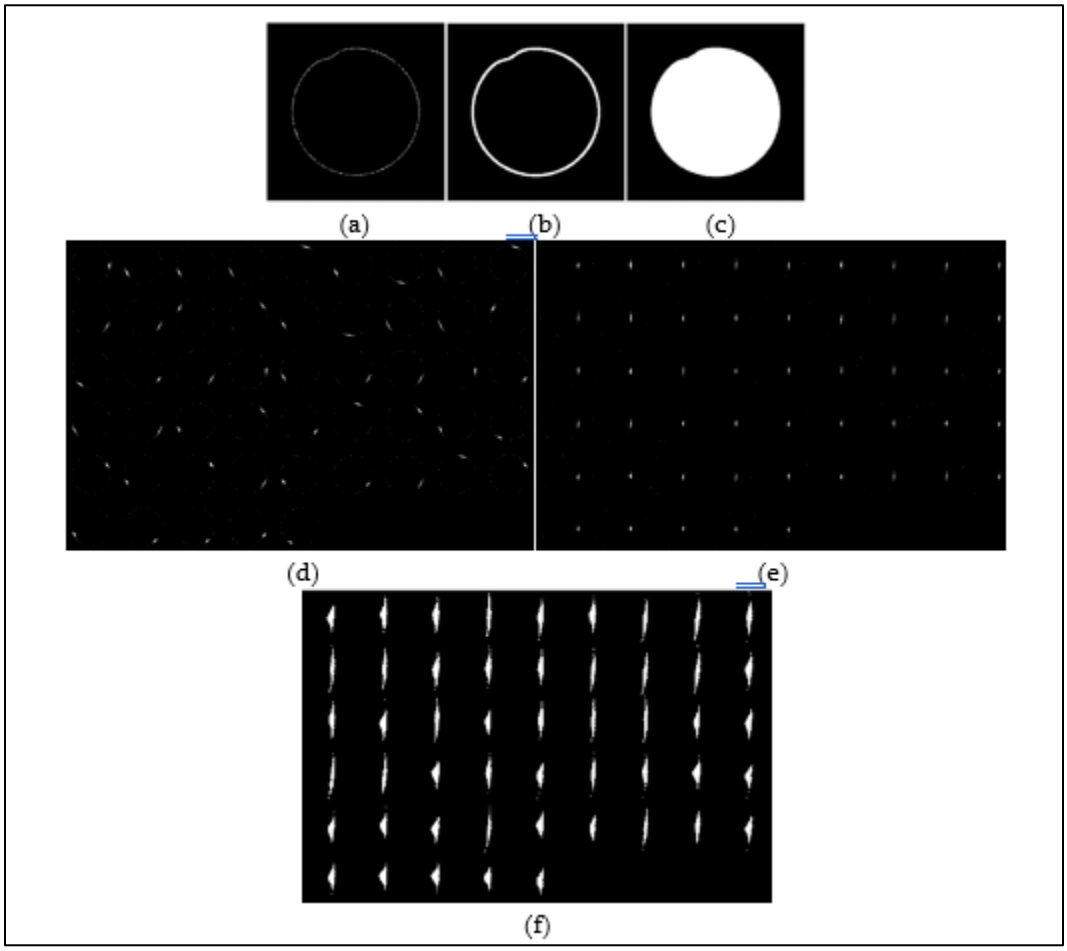


Figure 6-6. Image processing steps for extracting feature image. (a) the original frame for an impingement, (b) the dilation helps form a closed contour, (c) the closed circle is filled, (d) extracting the filled graphs from a round, (e) rotating the isolated damage into the same direction and (f) pooling to enlarge the damage and eliminate useless data

Using the feature image, the neural network can almost achieve no error in classification. It can be understood that a 2D image contains much more information than the simple 4 features data extracted from 3D model.

How would a Naïve Bayes behave with the same amount of information? A network with 3,600 nodes (one node for each pixel) were created in this case, each node corresponds to a pixel in the 60 by 60 feature images. Since the value of the pixels takes only 0 and 1, and we only need to differentiate the magnitude of probability, the probability distribution for each node is model as a Gaussian distribution with fixed variance. The average of the pixel value was taken as the mean and the variance was set to a certain value for each likelihood function. It is found that when choosing the variance to be 0.4, the network can achieve the

highest performance. It can be seen from Figure 6-7, with this formulation, the Naïve Bayes Classifier can also achieve an almost perfect result.

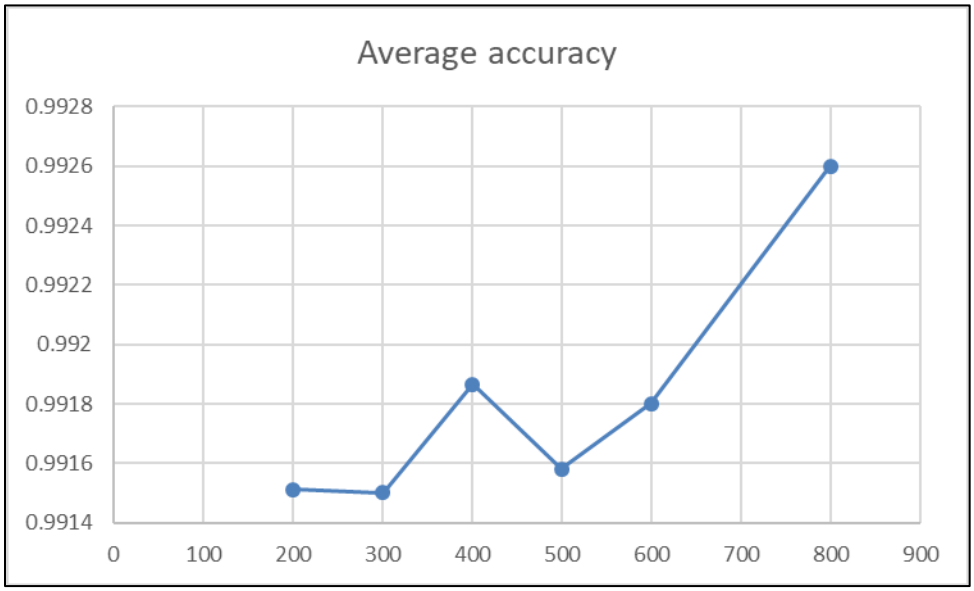


Figure 6-7. The accuracy vs. training data size using a refined Gaussian node

The above results indicated that a Naïve Bayes classifier is sufficient for the damage classification in our case.

Classification using real image

The real data was recorded using the prototype device received from the cooperating institute (Figure 6-8). It has a fisheye camera aligned with a laser ring projector. The two white plates are 3D printed and it can fit into a 3-inch pipe. The camera connects to a computer via USB port and the laser projector needs an additional power supply.

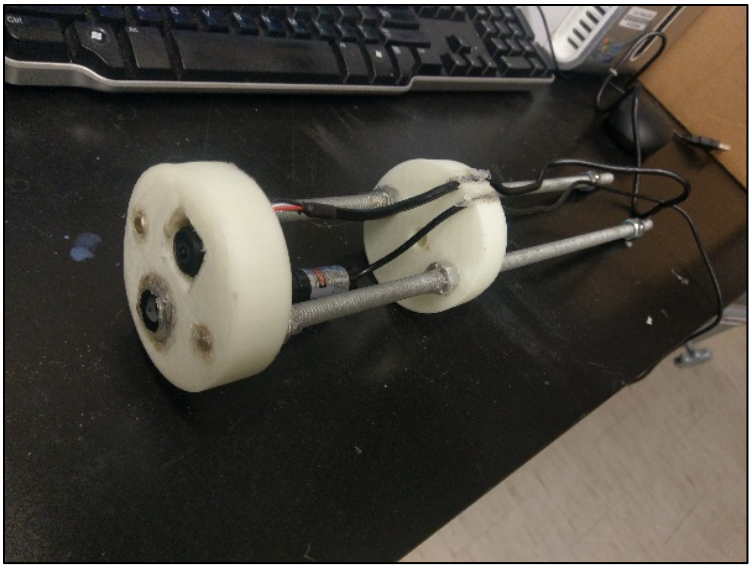


Figure 6-8. The prototype camera

To conduct the experiment, we bought a few PVC pipes from local hardware store and manually created/simulated some damage. **Figure 6-9** shows the manually created damages. A slit was cut using a saw on the pipe wall. Some pastes were stuck on the pipe wall to simulate dent and impingement.

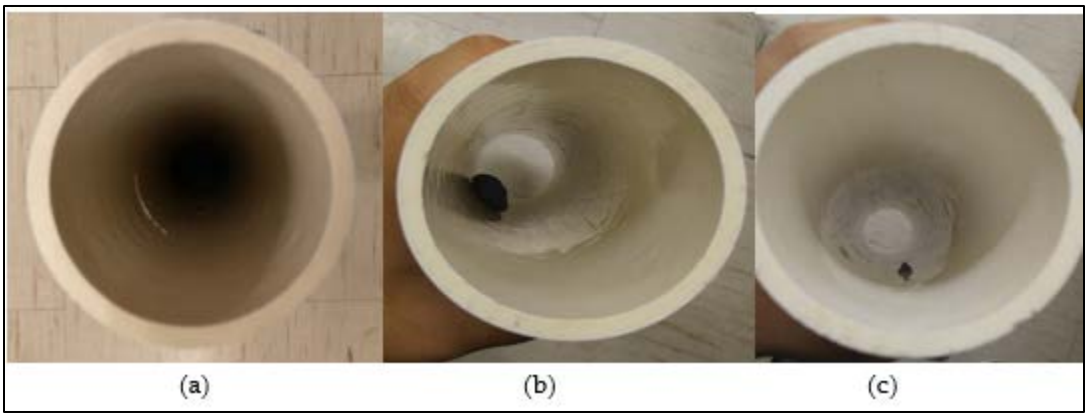


Figure 6-9. Damaged pipes with slit (left), dent (middle) and impingement (right)

But the reconstruction algorithm provided is very sensitive to the vibration of the camera. As we investigate the code, the algorithm simply stacks up all the black and white frames to form the 3D model. And this would lead to various bad visualization of the reconstructed pipe. To overcome these issues, photo shots, instead of videos, were taken when the camera is moving along the pipe. By doing this we could effectively reduce the vibration effect. Photo groups were taken with the damage set at different angle and paste for dent and impingement with different shape. The process of collecting data is a long and verbose process. After the data collection, each group of pictures were taken to the MATLAB

code to reconstruct the 3D pipe model. Some good results can be seen in Figure 6-10. And the 3D pipe model (essentially a 3D binary matrix) were processed using the denoise and frame averaging algorithm to isolate the damage. And based on the isolated damage, geometric features, such as length, volume, surface area, can be calculated.

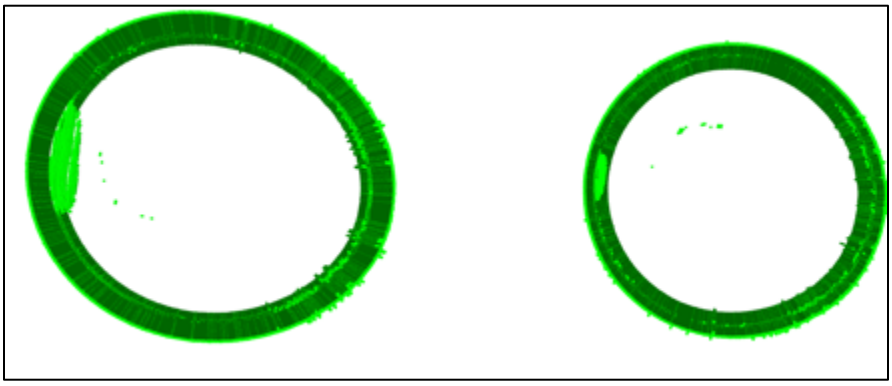


Figure 6-10. Two good reconstruction results

Due to the nature of the algorithm, out of all the tests only 90 sets of data, 30 sets for each kind of damage, can be selected to do classification. Naïve Bayes classifier were used to classify the data. By varying the size of the training data, the average accuracy is around 77%. Figure 6-11 shows the trend of average accuracy in relation to training data size.

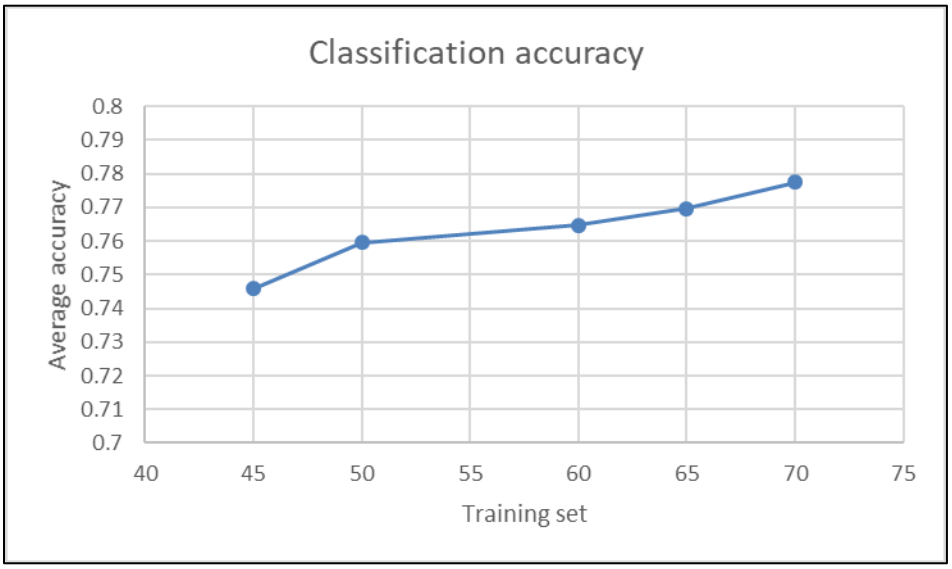


Figure 6-11. The average classification accuracy vs. training size using real data

Image reconstruction and damage classification conclusion

A simple Naïve Bayes classifier is applied in the damage detection and classification task. The method showed promising result with simulated pipe imaging. Its performance can be

comparative to more advanced method, such as Neural Network. This is probably due to that our case is not complicate in nature. The classification results using real imaging data was disappointing. This was mainly due to the prototype reconstruction device. The algorithm for the reconstruction needed to be revised.

Bayesian/maximum entropy network

The proposed maximum entropy (ME) network could introduce extra constraint on the features in a Bayesian network. In this section, we first introduce the concept of the maximum entropy. The detailed derivation of how the ME can be applied into a Bayesian classifier. A toy example is given to demonstrate the advantage of the proposed BEN network comparing to that of Naïve Bayes.

Introduction to maximum entropy

The Maximum relative Entropy (ME) method was originally used to assign probabilities using information as constraints [59]. Giffin [84] used the ME method to update probability with moment constraints. Caticha [61] found the relationship between the ME method and Bayes' theorem and showed that the Bayes' rule is a special case of ME. Guan [85] applied the ME updating into fatigue damage prognosis in a single updating scheme.

For an uncertain parameter $\theta \in \Theta$ and the corresponding response variable $x \in X$, let $\mu(x,\theta)$ be the prior of the joint probability distribution and $p(x,\theta)$ the posterior of the joint probability. According to maximum entropy axioms [86], the desired $p(x,\theta)$ would maximizes the relative entropy:

$$S[P, P_{\text{old}}] = -\int dx d\theta P(x,\theta) \log \frac{P(x,\theta)}{P_{\text{old}}(x,\theta)} \qquad \textbf{Equation 6-3}$$

The constraints with observation data *x'*, which is the case in a Bayesian updating scheme can be expressed as a delta function at *x'*:

$$P(x) = \int d\theta P(x,\theta) = \delta(x - x') \qquad \textbf{Equation 6-4}$$

And the moment constraint for a given function *f(θ)* is given as:

$$\int dx d\theta P(x,\theta) f(\theta) = F \qquad \textbf{Equation 6-5}$$

After forming a Lagrangian function including the constraint in Eq. (4) and (5), the updated function P is expressed in the form:

$$P_{\text{new}}(\theta) = P_{\text{old}}(\theta)\frac{P_{\text{old}}(x'|\theta)}{P_{\text{old}}(x')}\frac{e^{\beta f(\theta)}}{Z} \qquad \textbf{Equation 6-6}$$

where x' is the observation data and Z is a normalizing constant. The term β is determined by $\frac{\partial \log Z}{\partial \beta} = F$. This additional exponential term serve as a shifting factor from the old distribution. It is clear that the result when $\beta = 0$ will recover the Bayes' rule.

Maximum entropy in Bayesian network classifier

When applying the ME method into classification, the constraint information would be different than the above introductions. Recall a basic Naïve Bayes network as shown in **Figure 6-1** the class node C is a discrete node, the value of which is the class label. And the corresponding feature nodes f_1 to f_2 could be either continuous or discrete depending on specific setup. Each node contains its prior distribution and the edge contains the likelihood function between the two nodes. The constraint would be given on the likelihood function, since the prior knowledge would usually be the information about a feature given a certain class. For example, when distinguishing an orange to an apple, the information we know is that if it is an orange, it must have a rough surface and the color is orange. Detailed derivation will be shown as follows.

The constraint on the likelihood function given a known moment information is:

$$\int df_j\, p(f_j \mid C = c_i) g(f_j) = G_i \qquad \textbf{Equation 6-7}$$

(7)

where f_j is the jth feature in the network and c_i is the ith class label. The equation states that the expected value of some function $g(f_j)$ is G_i. The constraint is enforced on f_j corresponding to $C = c_i$. Additionally, we have the two normalization constraints for the likelihood function and the joint pdf:

$$\int df_j\, p(f_j \mid C = c_i) = 1 \qquad \textbf{Equation 6-8}$$
$$\int df_j dC p(f_j, C) = 1$$

Forming the Lagrangian function regarding the three constraint in Eq. (7) and (8) we have:

$$\begin{aligned}\mathcal{L} = &-\int df_j dC p(f_j, C)\log\frac{p(f_j, C)}{\mu(f_j, C)} + \alpha[\int df_j dC p(f_j, C) - 1] \\ &+\int dC\beta(C)[\int d f_j\, p(f_j, C) g(f_j) - G_i p(C)] \\ &+\int dC\gamma(C)[\int d f_j\, p(f_j, C) - p(C)]\end{aligned} \qquad \textbf{Equation 6-9}$$

α, $\beta(C)$ and $\gamma(C)$ are Lagrangian multipliers. Note that Eq. (7) and the first equation in Eq. (8) were multiplied with $p(C=c_i)$ on both sides and that $p(f_j, C=c_i) = p(f_j \mid C=c_i)\,p(C=c_i)$. Since the class label C is a discrete variable, the integral can be regarded as a summation. To find the optimal posterior $p(f_j, C)$, the variation of the Lagrangian function is se to be 0, i.e. $\delta\mathcal{L}=0$. This yields $\partial\mathcal{L}/\partial p = 0$:

$$\frac{\partial\mathcal{L}}{\partial p} = \int df_j dC[-\log\frac{p(f_j,C)}{\mu(f_j,C)} - 1 + \alpha + \beta(C)g(f_j) + \gamma(C)] = 0 \qquad \textbf{Equation 6-10}$$

The above equation satisfies for any $p(f_j, C)$ which means:

$$-\log\frac{p(f_j,C)}{\mu(f_j,C)} - 1 + \alpha + \beta(C)g(f_j) + \gamma(C) = 0 \qquad \textbf{Equation 6-11}$$

Hence:

$$p(f_j,C) = \mu(f_j,C)e^{-1+\alpha}e^{\beta(C)g(f_j)}e^{\gamma(C)} = \frac{\mu(f_j,C)e^{\beta(C)g(f_j)}e^{\gamma(C)}}{z} \qquad \textbf{Equation 6-12}$$

With

$$z = 1\big/e^{-1+\alpha} = \int df_j dC\mu(f_j,C)e^{\beta(C)g(f_j)+\gamma(C)} = \sum_C \int df_j \mu(f_j,C)e^{\beta(C)g(f_j)+\gamma(C)} \qquad \textbf{Equation 6-13}$$

By assuming the prior does not change, i.e. $p(C) = \mu(C)$, we can solve for the likelihood function:

$$p(f_j \mid C) = \frac{\mu(f_j \mid C)e^{\beta(C)g(f_j)+\gamma(C)}}{z} \qquad \textbf{Equation 6-14}$$

Back substitute into the constraints in Eq. (7) and the first equation in Eq. (8):

$$\frac{\int d f_j \mu(f_j \mid C)e^{\beta(C)g(f_j)}e^{\gamma(C)}g(f_j)}{z} = G_i \qquad \textbf{Equation 6-15}$$

$$\frac{\int d f_j \mu(f_j \mid C)e^{\beta(C)g(f_j)}e^{\gamma(C)}}{z} = 1 \qquad \textbf{Equation 6-16}$$

Given the above two equations, for a fixed value of $C = c_i$ we can solve for the corresponding unknown $\beta(C = c_i)$ and $\gamma(C = c_i)$. Since the term $e^{\gamma(C)} / z$ can be regarded as a normalizing constant for a fixed c_i, we can eliminate this term by the ratio of Eq. (15) and (16):

$$\frac{e^{\gamma(C)} \int d f_j \mu(f_j \mid C) e^{\beta(C) g(f_j)} g(f_j)}{e^{\gamma(C)} \int d f_j \mu(f_j \mid C) e^{\beta(C) g(f_j)}} = G$$

Equation 6-17

Now let us focus on a specific case where the constraint on likelihood function is a first order moment, i.e. $g(f_j) = f_j$. And the likelihood function is a normal distribution $\mu(f_j \mid C) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(f_j - \mu)^2}{2\sigma^2}}$. Substitute into Eq. (17) and eliminate the common term:

$$\frac{\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2} f^2 + (\frac{\mu}{\sigma^2} + \beta) f} f df}{\int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2} f^2 + (\frac{\mu}{\sigma^2} + \beta) f} df} = G_i$$

Equation 6-18

From basic calculus, we have: $\int_{-\infty}^{\infty} e^{-ax^2 - bx} dx = \sqrt{\frac{\pi}{a}} e^{b^2/4a}$

$$\int_{-\infty}^{\infty} x e^{-ax^2 - bx} dx = -\frac{\sqrt{\pi}}{2a^{3/2}} e^{b^2/4a}$$

Equation 6-19

Let: $a = \frac{1}{2\sigma^2}, b = -(\frac{\mu}{\sigma^2} + \beta)$

Substitute into (18):

$$\frac{-\frac{\sqrt{\pi} b}{2a^{3/2}} e^{\frac{b^2}{4a^2}}}{\sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a^2}}} = -\frac{b}{2a} = G_i$$

Equation 6-20

Solve for β:

$$\beta = \frac{G_i - \mu}{\sigma^2}$$

Equation 6-21

Substitute β into (16):

$$\frac{e^{\gamma(C)}}{z} = \frac{1}{\int d f_j \mu(f_j \mid C) e^{\beta(C) g(f_j)}}$$ **Equation 6-22**

Let:

$$I = \int d f_j \mu(f_j \mid C) e^{\beta(C) g(f_j)} = \frac{e^{-\frac{\mu^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2} f^2 + (\frac{\mu}{\sigma^2} + \beta) f} df$$ **Equation 6-23**

Again, we let: $$a = \frac{1}{2\sigma^2}, b = -(\frac{\mu}{\sigma^2} + \beta) = -\frac{G_i}{\sigma^2}$$

So:

$$I = \frac{e^{-\frac{\mu^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a^2}} = e^{\frac{G_i^2 - \mu^2}{2\sigma^2}}$$ **Equation 6-24**

Hence:

$$\frac{e^{\gamma(C)}}{z} = I^{-1} = e^{\frac{\mu^2 - G_i^2}{2\sigma^2}}$$ **Equation 6-25**

The new updated joint distribution can be written as:

$$p(f_j, C) = \mu(C) \mu(f_j \mid C) e^{\frac{G_i - \mu}{\sigma^2} f_j} e^{\frac{\mu^2 - G_i^2}{2\sigma^2}}$$ **Equation 6-26**

The result shows that after the updating with moment information, there would be 2 more exponential term. Which is similar to the result found in [84, 85].

Demonstration example for Bayesian/maximum entropy network

Using the above derivation, a maximum entropy network is formed and can be trained. Following is a simple example to illustrate the behavior of a BEN classifier against the NB classifier.

Assume a dataset of two types (classes) of damaged pipes, namely slit (class 1) and impingement (class 2). Two measurements were taken: the length is measured as the longest dimension for the damage and the volume of the damage. The data were randomly generated according to four independent normal distributions as listed in **Table 6-1**:

Table 6-1. Random generator for demonstration data

Damage type	Length	Volume

Slit	$N(175,5^2)$	$N(100,15^2)$
Impingement	$N(165,5^2)$	$N(150,20^2)$

A total of 200 data were randomly generated. Assume that we know by experience that the length of a slit is typically 100 unit. So the constraint can be mathematically expressed as:

$$\int df_2 \, p(f_2 \mid C = C_1) \, f_2 = 100 \qquad \textbf{Equation 6-27}$$

Substituting into the results from the above derivation, a BEN classifier is formed and can be trained. The classification with the example data can be seen in **Figure 6-12**. The comparison is the result from Naïve Bayes, i.e. no additional constraint. As we can see, due to the additional constraint information, the accuracy of the entropy network is significantly higher than that of a Naïve Bayes when there are less training data. When the training size increases, the accuracy from the two network converges.

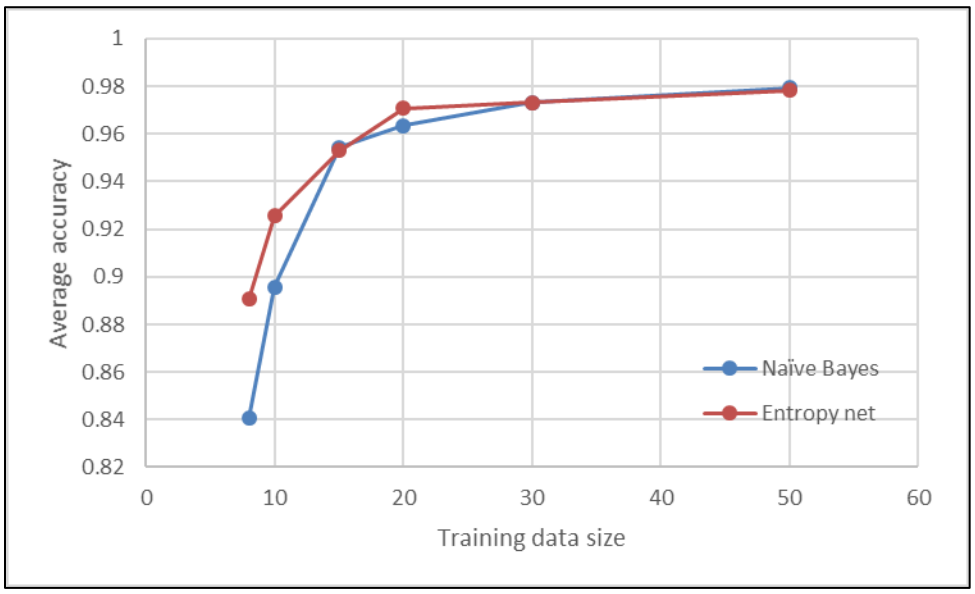


Figure 6-12. Average accuracy for NB and BEN

When we look at the probability distribution of the likelihood function for $p(f_2 \mid C = C_1)$ in both cases, in **Figure 6-13**, the additional constraint helps enforce that the mean of the variable is 100.

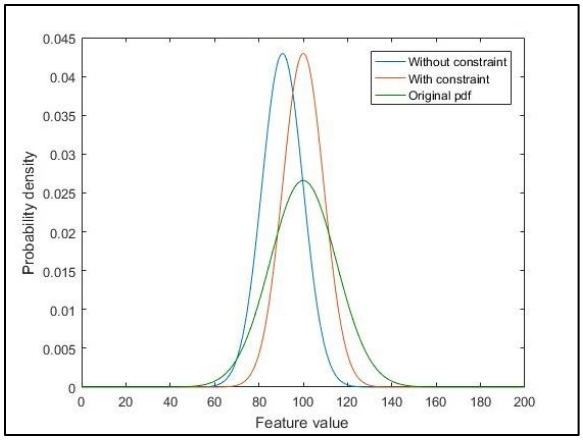


Figure 6-13. The comparison of the updated probability distribution with and without constraint

It can be concluded that the entropy network does improved classification accuracy when training data is small. Which is often the case for some engineering problems. Such a method could benefit classification tasks where experimental data is hard to get.

Bayesian/maximum entropy network conclusion

In this section, the maximum entropy method was applied in the Bayesian network and used as a classifier. Detailed derivation showed that it was possible to accept any order of moment constraint. An analytical solution was given for a special case where the nodes are modeled as Gaussian and the constraint as a first order moment. Theoretically, any constraint that could be expressed in the form of an equality of the probability function could be encoded into the classifier. A toy example was given to show that the BEN indeed performs better when the training data is small. Since the empirical information on the pipe imaging data is not available, the method was not tested against real data.

Creep crack growth prediction

In this section, an equivalent crack growth model for creep life prediction of polymers is discussed. The equivalent crack growth rate is modeled using a Paris' law like equation. The model was calibrated and validated using the data provided by GTI. The experimental data was done on Aldyl-A pipes under various loading condition and different damage types. This enabled the study of the effect of damage on the life of the pipe materials. The damage is considered as stress risers in the pipe. The results showed that the method has a good prediction comparing to the experimental data. The prediction results was then applied to the maintenance framework.

Model development

The proposed method expressed the crack growth rate as a Paris' type function as:

$$\frac{da}{dt} = C \cdot K^m$$ **Equation 6-28**

where C and m are material properties that will be calibrated using the control data group. K is the stress intensity factor (SIF). The crack is assumed to be a semi-circular crack at the inner surface in the longitudinal direction. The SIF solution is given in [65] as:

$$K = \sigma \sqrt{\frac{\pi}{Q} a} F$$ **Equation 6-29**

where a is the crack length, σ is the hoop stress, Q is the shape factor, in this case is given a fixed value of Q=2.464 and F is the boundary correction factor. An asymptotic SIF solution considering the notch effect is proposed in [66]. The expression for the SIF is modified as:

$$K = \sigma \sqrt{\frac{\pi}{Q}(a + d\{1 - \exp[-\frac{a}{d}(K_t^2 - 1)]\})} F$$ **Equation 6-30**

where K_t is the stress concentration factor, d is a geometric measurement of the notch size. A schematic illustration is shown in Figure 6-14. t is the thickness of the pipe. The boundary correction factor is fitted as a function that is related to the geometry of the crack. In this specific case, F is calculated as [65]:

$$F = (1.04 + 0.2017 \times (\frac{a}{t})^2 - 0.1061 \times (\frac{a}{t})^4) \times (6.05 - 0.5\sqrt{\frac{a}{t}})$$ **Equation 6-31**

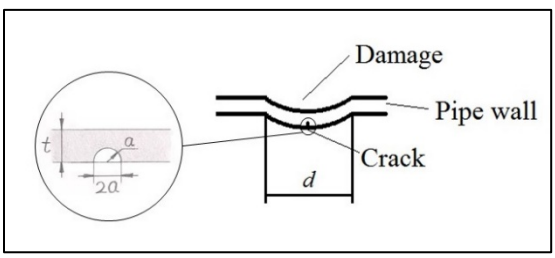


Figure 6-14. Schematic plot of the crack at a notch root and the semi-circular geometry of a crack

Integrating the crack growth rate equation from initial crack length a_i to critical crack length a_c, the failure time T_f can be expressed as a function of the hoop stress σ as

$$T_f = C^{-1} \int_{a_i}^{a_c} K^{-m} da = C^{-1} (\sqrt{\frac{\pi}{Q}})^{-m} \sigma^{-m} \int_{a_i}^{a_c} F \sqrt{a + d\{1 - \exp[-\frac{a}{d}(K_t^2 - 1)]\}}$$ **Equation 6-32**

Taking the logarithm on both side of the equation gives a linear relationship between the failure time and stress as:

$$\log T_f = -m * \log \sigma + \log(C^{-1}(\sqrt{\frac{\pi}{Q}})^{-m} I)$$ **Equation 6-33**

where $I = \int_{a_i}^{a_c} F\sqrt{a + d\{1 - \exp[-\frac{a}{d}(K_t^2 - 1)]\}}da$. Eq. (33) is the proposed life prediction model for the slow crack growth (SCG) of polymers.

To predict the crack growth as a function of time, the upper limit of the integral in I can be changed to an arbitrary crack length a_t. And the corresponding time on the left-hand side can be calculated. This gives an implicit equation for the crack length as a function of time and stress level with five model parameters. It can be solved by numerical method and the crack length vs. time can be plotted.

$$t = C^{-1}(\sqrt{\frac{\pi}{Q}})^{-m} \sigma^{-m} \int_{a_i}^{a_t} F\sqrt{a + d\{1 - \exp[-\frac{a}{d}(K_t^2 - 1)]\}}da$$ **Equation 6-34**

Calibration and validation

The data set from GTI was done using Aldyl-A pipe. It consists of 5 groups, namely the ductile control group, SCG (slow crack growth) control group, LDIW (low ductile inner wall) control group, LDIW with indentation and LDIW with squeeze-off. The ductile group is tested in high stress level and failure is dominated by ductile deformation. There is a total of 450 data points in this data set. The hoop stress and the failure time were recorded and shifted to 23°C (73.4°F), which is the normal operating temperature. The test used Rate Process Method (RPM) to evaluate the long-term performance of the pipe and used bi-directional shift factors to shift the test results to a common reference temperature. The properties of the polymeric material are governed by the activation energy of the molecular rearrangement. It is measured using Dynamic Thermo-Mechanical Analysis techniques. The activation energy can help determining the Time Temperature Superposition characteristics of the material [40][2], which are the basis of the RPM calculations. GTI employed DTMA to measure the activation energy and use the measurement to develop the bi-directional shift factors [8]. By using these information, the test results at an elevated temperature can be shifted to the reference temperature. The stress concentration factors in the data set are estimated by dividing the actual shifted time by the mean time calculated from control SCG model. Since the study focuses on the creep behavior, the ductile control group will not be used. The SCG data will be used to calibrate the material properties in the proposed model, and the other groups will be used to validate the prediction result of the model. The data are

plotted in **Figure 6-15** in double log scale. The log-log plot for the stress vs. life of the data showed a linear tendency.

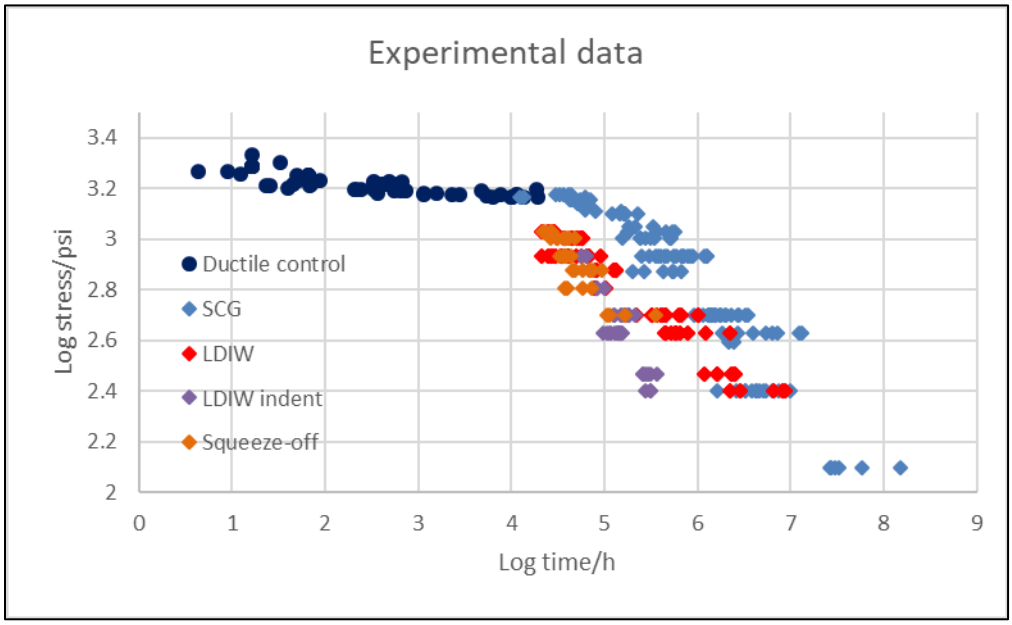


Figure 6-15. Experimental data shifted to 23°C (73.4°F)

Figure 6-16 shows the geometry of the pipe. The outer diameter is 2.375 in. And the SDR value is 11, which means the ratio of outer diameter and pipe thickness is 11.

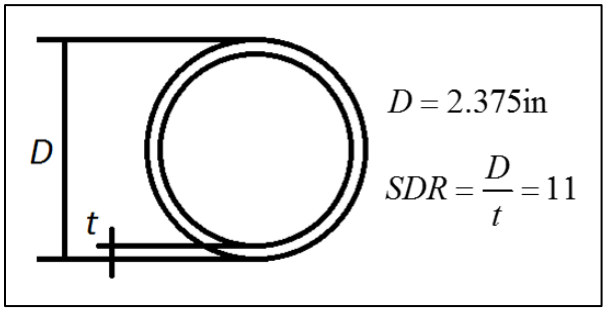


Figure 6-16. The geometry of the pipe

The initial crack length can be observed through SEM image in **Figure 6-4,** the micro-crack measures about 25 μm. In this case, the initial crack length is set as $a_i = 10^{-3}$ in . The critical crack length is assumed to be $a_c = 0.1$ in . The measurement for the critical crack length is not important because the crack growth rate will be extremely fast when closing to rupture. A typical damage size is around 1 inch, the value for d is chosen to be $d = 1$ in .

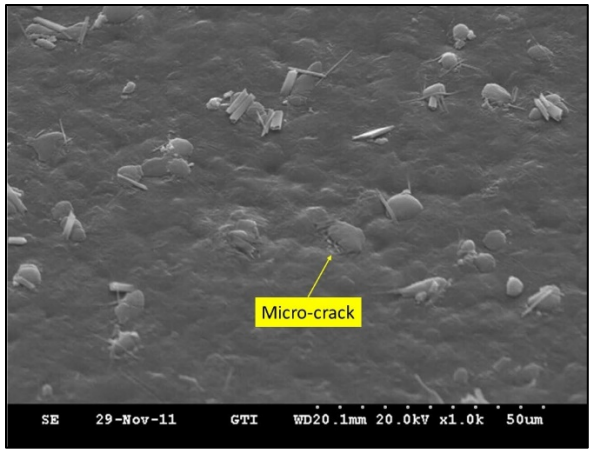


Figure 6-17. The SEM image showing an initial crack

The data in SCG group is used as reference data, a linear relation between the logarithm of stress and time is calculated by linear regression. The linear regression gives the stress as a function of time as: $\log T_f = -2.5632 \log \sigma + 13.0796$, indicating $m = 2.5632$. The mean of the stress concentration factor is used as the value for K_t to evaluate the integral I. And the value for C is calculated as: $\log C = -13.3442$. Here, the value of C is considered as a random variable. Assume that the slope of the log linear curve, m, is fixed. By back substituting the experimental data points into the equation, we could get an array of $\log C$. The value for the logarithm of C can be fitted into a normal distribution. For the SCG data, $\log C$ follows a normal distribution with mean -13.3442 and variance 0.2960. Using these data, we could plot the regressed stress vs. life curve for SCG data in **Figure 6-18**.

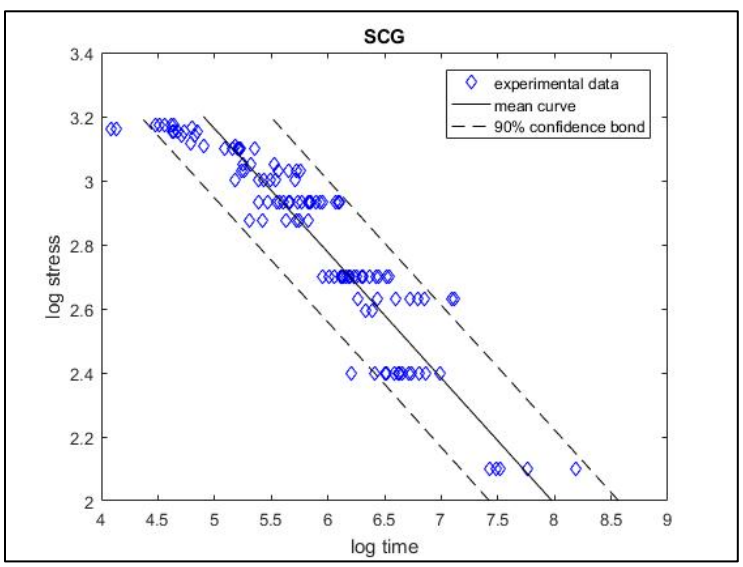


Figure 6-18. Stress-Life curve for SCG data and its confidence bond against the lab data

By substituting the K_t with the mean value from other groups, the stress vs. life can be predicted for other damage types. The results of the prediction versus the experimental data for LDIW, indentation and squeeze-off are plotted in **Figure 6-19**. The predictions agree with the data closely.

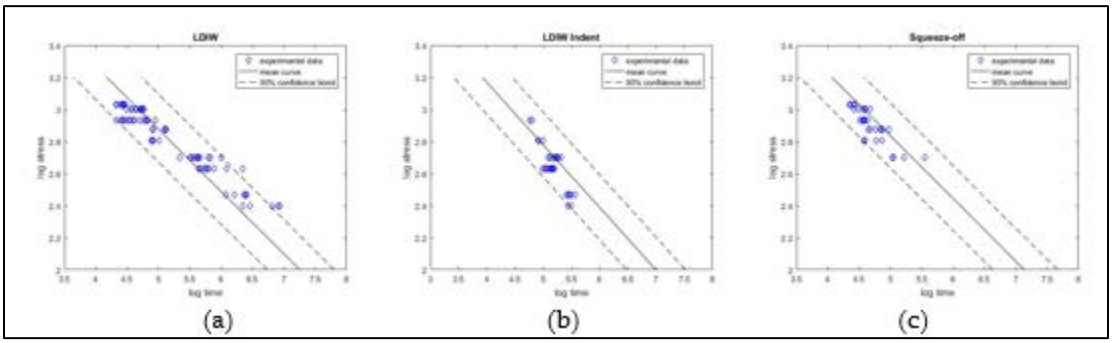


Figure 6-19. The prediction of stress-life curve for (a) LDIW, (b) LDIW indent and (c) Squeeze-off

Uncertainty and reliability prediction

The value of C and Kt in the prediction model can be regarded as random variables. The distribution of $\log C$ can be found by fixing m and substitute the experimental data and is calculated above to follow a normal distribution with mean -13.3442 and variance 0.2960. The stress concentration factor K_t in each group was tested against various types of distribution by KS test [87] and is found that a logistic distribution can best describe the measured data. The uncertainty quantification of the model parameter for SCG group is listed in **Table 6-2**.

Table 6-2. Uncertainty quantification of the model parameters for SCG group

Parameter name	Distribution	Distribution parameter	
		Mu	sigma
a_i	Constant	0.001	
a_c	Constant	0.1	
d	Constant	1	
$\log C$	Normal	-13.3442	0.2960
m	Constant	2.5632	
K_t	Logistic	2.6192	0.2445

The crack growth prediction is achieved by numerically evaluating the integral I at continuous values of crack length a_t given a pair of random sample of C and K_t. Hence, the implicit function of a_t as a function of time can be plotted. **Figure 6-20** shows the crack growth curve for SCG at $\sigma = 1000$ psi along the 90% confidence bond. It shows that the

crack growth process is quite slow in the beginning and gradually speeds up. This plot could be used to calculate the failure probability at a given time.

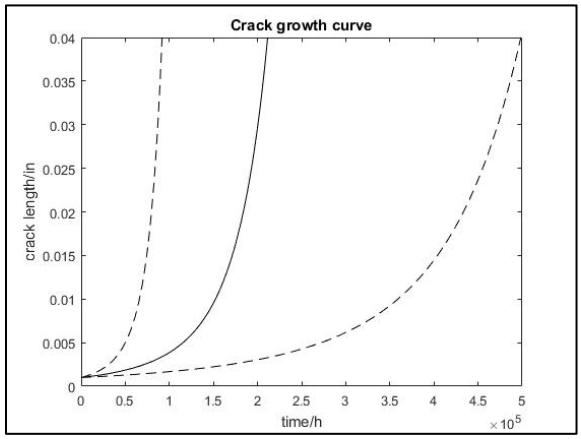


Figure 6-20. The crack growth curve for SCG and its confidence bond at 1000 psi

The probability of failure is studied at $\sigma = 1000$ psi and $t = 2\times10^5$ h. The mean curve in Figure 6-20 indicates that the crack length at $t = 2\times10^5$ h is around 0.03 inch, which is still far from failure. While the MC method simulation of crack length at $t = 2\times10^5$ h indicates that it is possible for the crack length to exceed the critical crack length. Figure 6-21 shows the empirical cumulative density function (CDF) and probability density function (PDF) generated from 10,000 MC samples. As we can see in the CDF plot, only a little over 60% of the simulated MC sample is beneath the critical crack length. The shaded area in **Figure 6-21** (b) gives the failure probability. For a deterministic model, a crack length of 0.03 inch would not be considered as failure. But according to the probability analysis, the failure probability at $t = 2\times10^5$ h is 37.85%, which is not negligible. The failure probability model would be significant in reliability analysis.

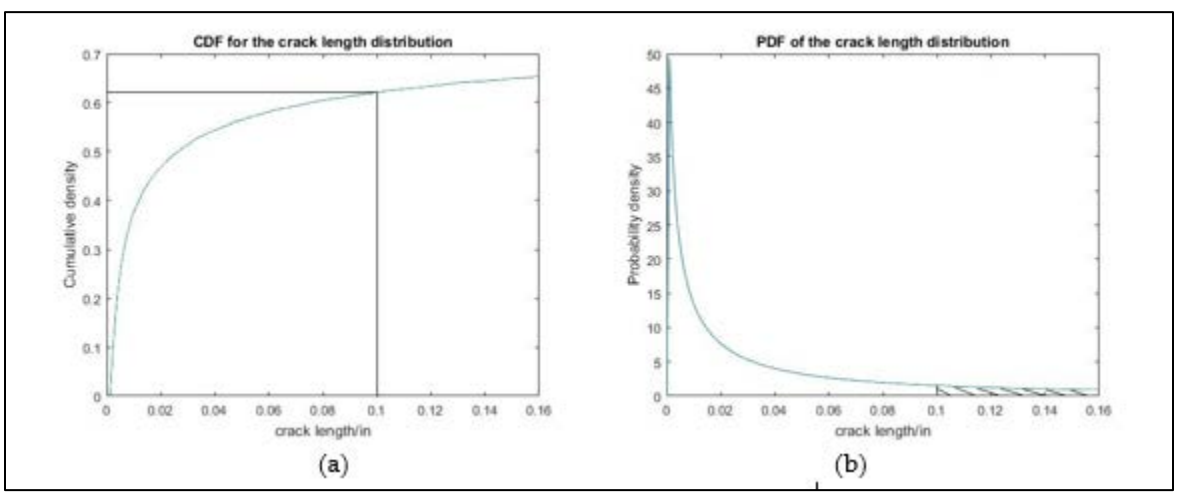


Figure 6-21. The empirical CDF and (b) PDF of the crack length distribution

Approximated creep crack growth prediction conclusion

The prediction model used a simple power law equation to predict the crack growth in polymer pipes. The novelty of the method is that it considered the effect of damage on the life prediction. The error analysis was done by taking the difference in failure time of the lab data against the predicted mean. For each data points, the error is calculated as the actual life from test minus the predicted mean failure time at the corresponding stress. The errors were collected and plotted as an empirical CDF in **Figure 6-22** compared with the MC sample errors. The MC sample error represents a standard distribution of the failure time.

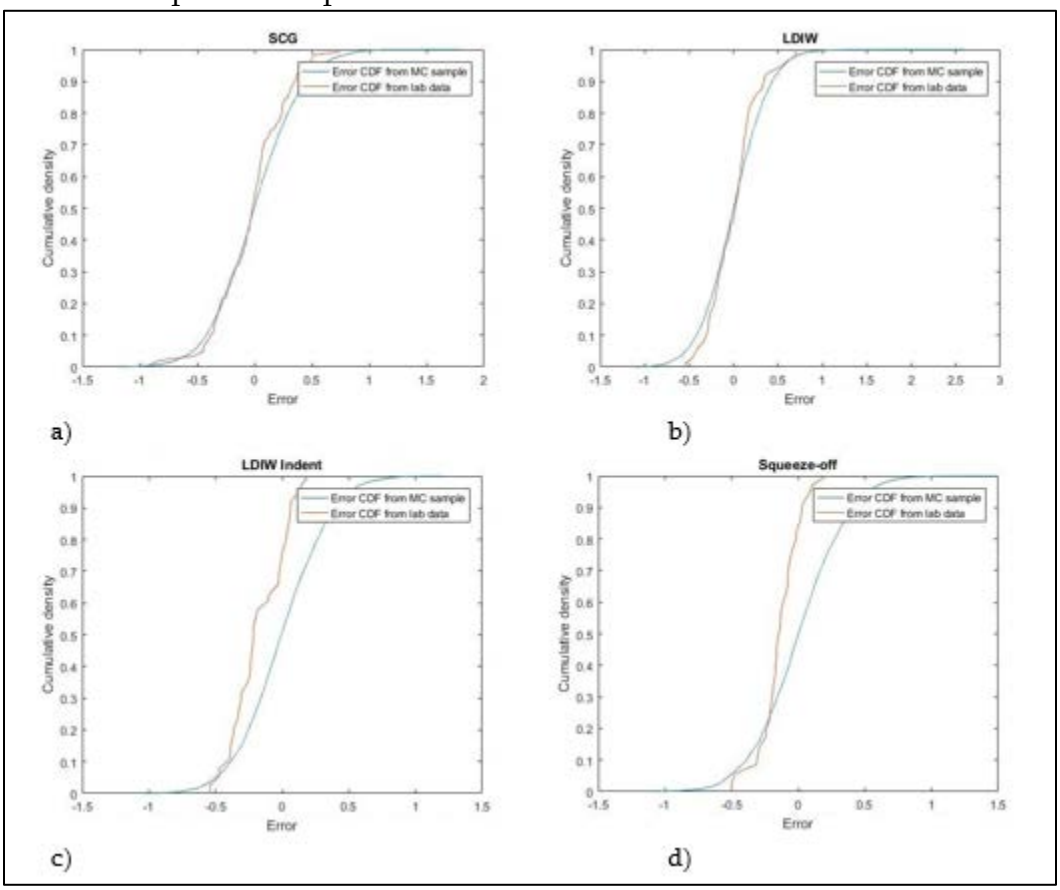


Figure 6-22. The empirical CDF of the error comparing with the CDF of the MC sample error for a) SCG, b) LDIW, c) LDIW indent and d) Squeeze-off

Within the margin of error, the experimental data agrees well with the predicted data from the proposed model. The group with indentation and squeeze-off deviates from the MC data, one possible explanation is that the assumed damage size d is different regarding these two types of damage.

Maintenance framework

In this section, a condition-based maintenance framework was formulated. The method categorizes the pipes into different condition stages. Based on the current condition and the

future working condition of the pipe, the condition vector and the cost for maintenance can be iteratively calculated. Minimizing the overall cost would give the optimal maintenance plan.

Maintenance framework formulation

Table 6-3 the list of symbols used in the formulation.

Table 6-3. Nomenclature for the maintenance decision framework

Name	Meaning
Q	The total quantity of pipes
S	Number of deterioration stage
M	Number of possible maintenance method
D	Condition vector with dimension 1×S
P	Deterioration matrix, S×S
$\mathbf{M}_m$	Maintenance transition matrix, S×S
X	Decision matrix, M×S
C	Cost matrix, M×S

Each term in the condition vector D represents the percentage of samples in each stage. The elements in the vector should sum up to 1. The degradation matrix is the probability transition matrix for a certain time step Δt. The term **P**(i,j) means the probability of transition from condition i to condition j. For example if **P**(2,5)=0.1, it means a sample that is now in condition 2 has a 10% probability of transiting to stage 5 after time period Δt, without any repair. The elements in each row of the degradation matrix should sum up to 1.

The maintenance matrix is the probability transition matrix for a maintenance method. The term $\mathbf{M}_m$(i,j) means the percentage of pipes that transit from condition i to condition j right after the maintenance method m. For example, $\mathbf{M}_2$(3,1)=0.1 means that 10% of the pipes that is in condition stage 3 will transit to condition 1 after maintenance method 2. The elements in each row should sum up to 1. The maintenance matrix for doing no maintenance is an identity matrix.

The decision matrix is the object of the optimization. The term **X**(i,j) means the percentage of pipes that is in condition j has maintenance i done. For example, **X**(2,4)=0.1 means that maintenance method 2 is applied to 10% of the pipes in condition stage 3. The elements in each column should sum up to 1.

The element **C**(i,j) in the cost matrix corresponds to the expense of applying a maintenance i for a pipe that is in condition stage j. For example, the element **C**(3,5)=5000 means it takes $5,000 to do a maintenance 3 to a pipe that is in condition stage 5.

When calculating the new condition vector after a time period Δt, the following equation is used:

$$\mathbf{D}_{new} = \sum_{m} \mathbf{D} \cdot\times \mathbf{X}(m,:) \times \mathbf{M}_{m} \times \mathbf{P} \qquad \textbf{Equation 6-35}$$

where the term $\mathbf{X}(m,:)$ is the mth row of the decision matrix, meaning the decision vector for maintenance m. The sign $\cdot\times$ is an element wise operator. The dot product of **D** and **X**(m,:) gives a $1\times S$ vector, which means the percentage of samples in each condition that will have maintenance m applied. The vector is then multiplied by the maintenance matrix, which would result in another $1\times S$ vector, means the condition vector right after maintenance m being applied. The condition after maintenance times the degradation matrix gives us the new predicted condition vector ($1\times S$) for the group of pipes that had maintenance m done. The sum over m adds up all the condition vector for different maintenance group.

The corresponding cost is calculated by:

$$\text{Budget} = \sum_{m} Q \times \mathbf{D} \cdot\times \mathbf{X}(m,:) \times \mathbf{C}(m,:) \qquad \textbf{Equation 6-36}$$

The dot product of **D** and **X**(m,:) times Q gives a quantity vector with dimension $1\times S$. Each value means the number of pipes in the corresponding condition that had maintenance m done. The quantity vector times the cost vector, with is the mth row of the cost matrix, yields the cost for doing maintenance m. The sum over all maintenance method gives us the total cost.

Generating probability transition matrix

In the previous section, we have predicted the approximated creep crack growth curve in a probabilistic sense, as in **Figure 6-20**. In this section, we will calculate the probability transition matrix for the maintenance frame work based on the crack growth curve.

In the following example, we are getting the probability transition matrix for a time steo of Δt=1500 h and for the stress level of $\Delta\sigma = 1000\text{psi}$. The deterioration stage is categorized into 5 stages from initial crack length to failure, listed in **Table 6-4**. We assume that the observation of the crack length is a uniform distribution, which means the observation of a crack length a is equally possible in the range of 0.001 to 0.1 in.

Table 6-4. The condition stage

Condition	Excellent	Good	Fair	Poor	Failure
Crack length/in	$10^{-3} – 2*10^{-3}$	$10^{-3} – 6*10^{-3}$	$6*10^{-3} – 2*10^{-2}$	$2*10^{-2=}$ -- 0.1	0.1<

Monte Carlo simulation were used in this case. We evenly sampled around 2000 points from 0.001 to 0.1 in as the observed crack length in current stage. Each sample point has a corresponding pair of *C* and *Kt* from the creep crack growth model. And the life after Δt is calculated by letting the current point evolve along the specified curve. **Figure 6-23** may help better understand the process of the MC simulation. As we can see in the picture, for an arbitrary initial crack observation a_i, we can find the corresponding life t_1 from the red curve. The crack length after time step Δt, regarding the same red curve, would be the crack length a_t corresponding to the life of t_1+Δt on the red curve. The blue curve is another possible curve with a different pair of *C* and *Kt*, and there is a corresponding life t_1' for the initial observation. The prediction of the crack length after time step Δt, according to the blue curve, is a_t' as shown. As we can see that the initial observation falls in condition stage 2, and the prediction according to the red curve would be in condition 3, while the prediction according to the blue curve would be in condition 4. Thus, for this initial observation, the term **P**(2,3)=0.5 and **P**(2,4)=0.5 since they have equal chances of transitioning to condition 3 and 4 after time Δt.

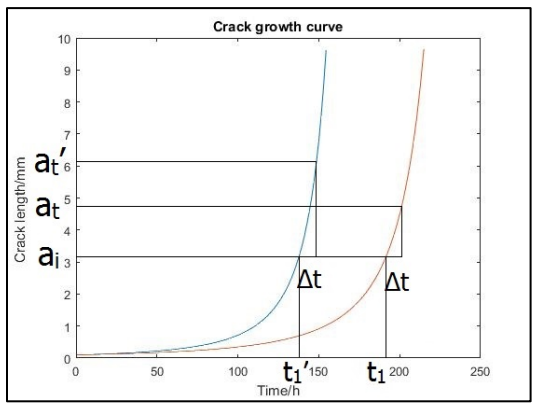


Figure 6-23. Illustration of the crack growth evolvement.

This process was done for all the curves in the cluster and the 2000 samples of observation. And the elements in the degradation matrix were calculated as the percentage of pipes that falls in each condition. For this example, the resulting degradation matrix is:

$$\mathbf{P} = \begin{bmatrix} 0.7945 & 0.2051 & 0.0003 & 0.0001 & 0 \\ 0 & 0.8121 & 0.1875 & 0.0003 & 0 \\ 0 & 0 & 0.8085 & 0.1914 & 9.34\text{e-}05 \\ 0 & 0 & 0 & 0.8999 & 0.1001 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \textbf{Equation 6-37}$$

Illustrative example

This section will go over the maintenance optimization problem through a demonstrative example, hoping that the reader can have a clear understanding of the proposed framework.

Maintenance decision optimization

As a demonstration example, we assume three type of maintenance: do nothing (a natural degradation process), repair and replacement. Still, we consider the pipes as in SCG condition and under a constant pressure of 1000 psi. The maintenance transition matrix for doing nothing is an identity and the one for repair and replace were assumed:

$$\mathbf{M}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.5000 & 0.5000 & 0 & 0 & 0 \\ 0.1000 & 0.3000 & 0.6000 & 0 & 0 \\ 0.0100 & 0.0400 & 0.2500 & 0.7000 & 0 \\ 0 & 0.0100 & 0.0400 & 0.1500 & 0.8000 \end{bmatrix}$$

$$\mathbf{M}_3 = \begin{bmatrix} 0.99 & 0.01 & 0 & 0 & 0 \\ 0.99 & 0.01 & 0 & 0 & 0 \\ 0.99 & 0.01 & 0 & 0 & 0 \\ 0.99 & 0.01 & 0 & 0 & 0 \\ 0.99 & 0.01 & 0 & 0 & 0 \end{bmatrix}$$

Equation 6-38

respectively. The cost matrix is assumed as:

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 30 & 30 & 50 & 100 & 200 \\ 1600 & 1600 & 1600 & 2500 & 3000 \end{bmatrix}$$

Equation 6-39

The initial conditions were set as:

$$\mathbf{D} = \begin{bmatrix} 0.10 & 0.20 & 0.50 & 0.15 & 0.05 \end{bmatrix}$$

Equation 6-40

The problem is that we are trying to minimize the cost under the reliability constraint: The failure probability should not exceed 0.05. The failure probability is the last term in the condition vector. The problem is transformed as to minimize the budget given the constraint of the last element in **D** should be less than 0.05. Due to the property of the maintenance decision matrix, the number of unknowns are 10.

Since the problem has 10 degree of freedom (DOF), genetic algorithm (GA) [88] was chosen to calculate the optimal maintenance planning. GA is inspired by bio-processes such as mutation and gene selection, and is commonly used in optimization problems. The method has advantages when dealing with high dimensional problems. Since GA is a non-

deterministic algorithm, the result for every run would be different. And the solution may not be the best solution. It is an efficient algorithm and has a MATLAB built in function.

The objective function for the maintenance problem is the total cost for maintenance as in **Equation 6-36**. The constraint is the last term in the condition vector.

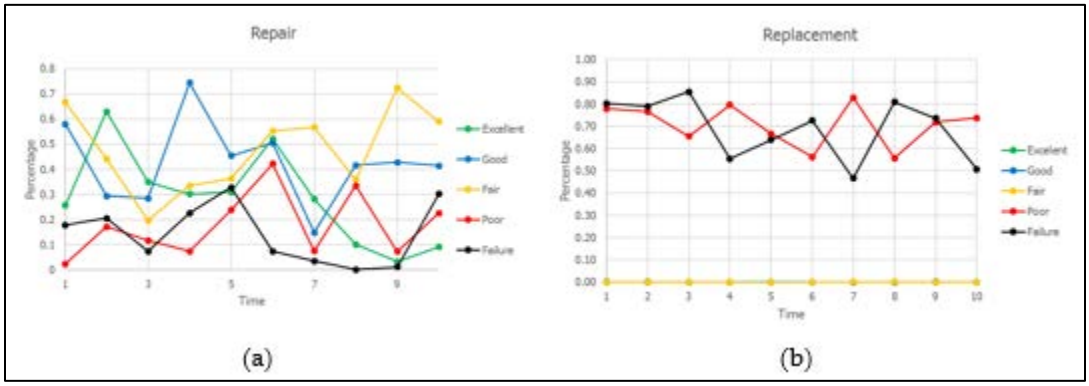


Figure 6-24. **Figure 6-24** shows the optimized results for 10 consecutive time steps. As we can see from the figure, more effort will be put in the replacements of pipes in poor and failure conditions and repair with the ones in good and fair conditions. Note that the result for the maintenance decision would differ from different runs, but the overall trend would be the same.

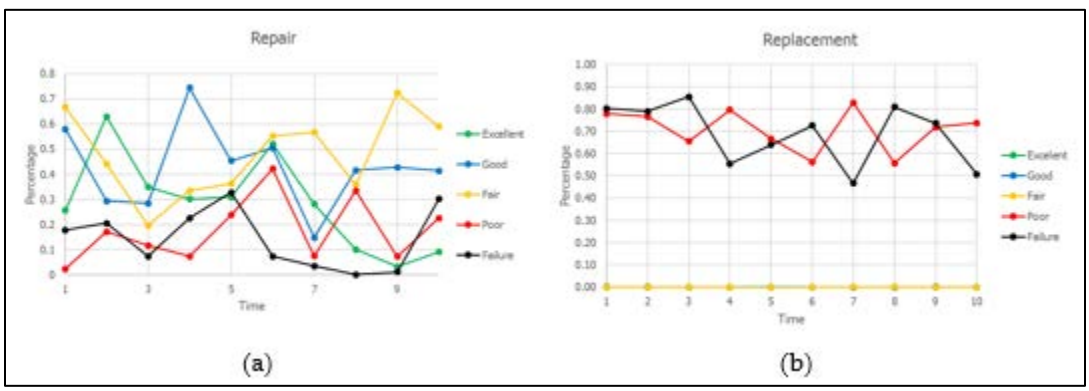


Figure 6-24. The visualization of maintenance plan for (a) repair and (b) replacement

Consequence cost

In general situations, the damage caused by failure is more severe in some cases while in other cases not. For example, an explosion caused by failure of gas pipe would be hazardous if it happened in a hospital, but not so if it were a rural area. So, per different circumstances, the importance of failure might be different. In regard of this, we add the consequence cost and different groups into the maintenance framework.

The idea is as simple as adding an extra dimension regarding the groups. We use G to represent the total number of groups. The condition vector now becomes a condition matrix with dimension $G \times S$. The quantity becomes a quantity vector of size $G \times 1$ with each term representing the quantity of pipes in each group. An additional weighting vector **W** of dimension $1 \times G$ was created, which is the weight (the severity of consequence) assigned to different group. The reliability objective function is:

$$reliability = \mathbf{W} \cdot \mathbf{D}_{new}(:, S) \qquad \textbf{Equation 6-41}$$

The term $\mathbf{D}_{new}(:, S)$ is the last column of the new predictive condition matrix, which represents the failure probability in each group. The product gives a scaler value.

Following the same example in the previous section, we assumed 3 groups of pipes (G=3), each represents a business center area, residential area and rural area. The three groups have the same initial observation of the condition vector as in the previous example. The quantity vector $\mathbf{Q} = [100 \quad 100 \quad 100]$, meaning 100 pipes in each group. We assume that the consequence cost of the three area is from high to low, and the weighting vector $\mathbf{W} = [10 \quad 5 \quad 1]$. The decision matrix now becomes a 3-dimensional matrix ($M \times S \times G$). The number of DOF increased to ((M-1)$\times S \times G$=) 30.

Figure 6-6 shows the total cost for each group. We can see that there will be more money spent in the group with higher consequence cost to reduce the probability of failure. This makes sense since if we wanted to avoid failure in one area it is reasonable to spend more money in maintain the pipes are in good working conditions.

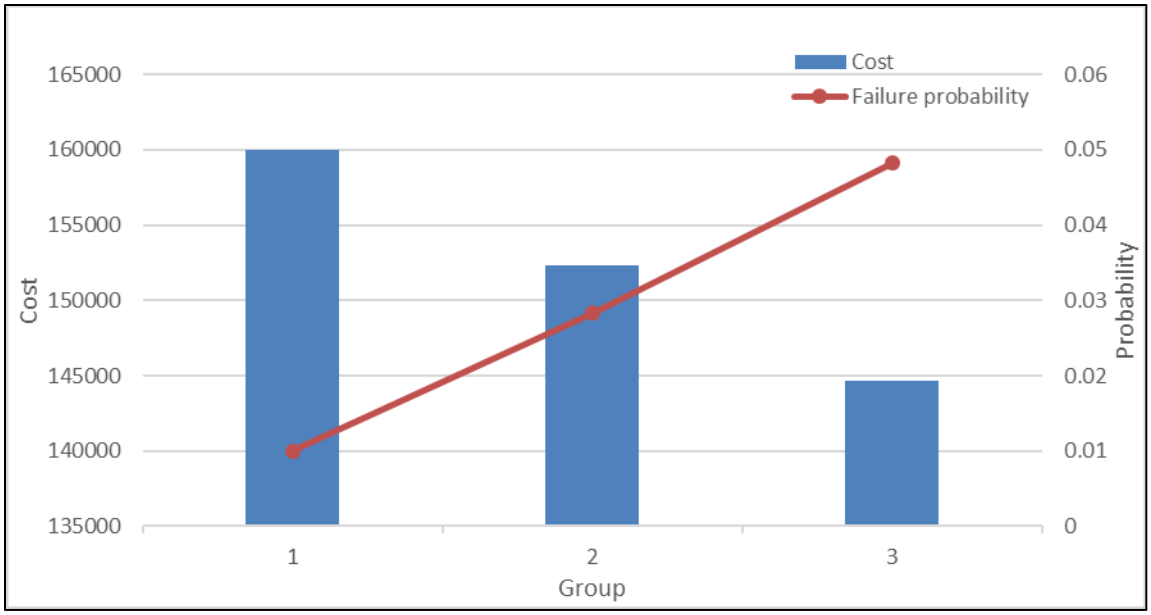


Figure 6-25. The cost and failure probability prediction for each group

Dynamic maintenance framework

The dynamic updating of the transition matrix can increase the prediction accuracy of the crack length. Hence help the maintenance planning to avoid unwanted failure or the unnecessary cost. The updating is achieved through the observation of *Kt* parameters. With a calibrated camera that reconstruct the pipe in 3D, the actual dimension of the damage can be measured. The stress concentration factor can hence be calculated using finite element model.

Bayesian updating is widely used in probability related studies. It is used to update the belief of existing knowledge (prior) through new information (observation). The updated posterior of a Bayesian updating network is:

$$P(\theta \mid x') \propto P(x' \mid \theta) P(\theta) \qquad \textbf{Equation 6-42}$$

P(θ) is the prior, which, in our case, is the distribution of *Kt* fitted by the lab data. P(x'| θ) is the likelihood function through observation.

Suppose a few observations of the Kt were made through the field data. And is used to update the Kt distribution. The observation is assumed to follow a normal distribution with mean of 4.2 and variance of 0.35. Markov Chain Monte Carlo (MCMC) method is used to generate the updated distribution samples. **Figure 6-26** shows the updated probability compared with the original one. We can see that the updated value of *Kt* has smaller mean and variance.

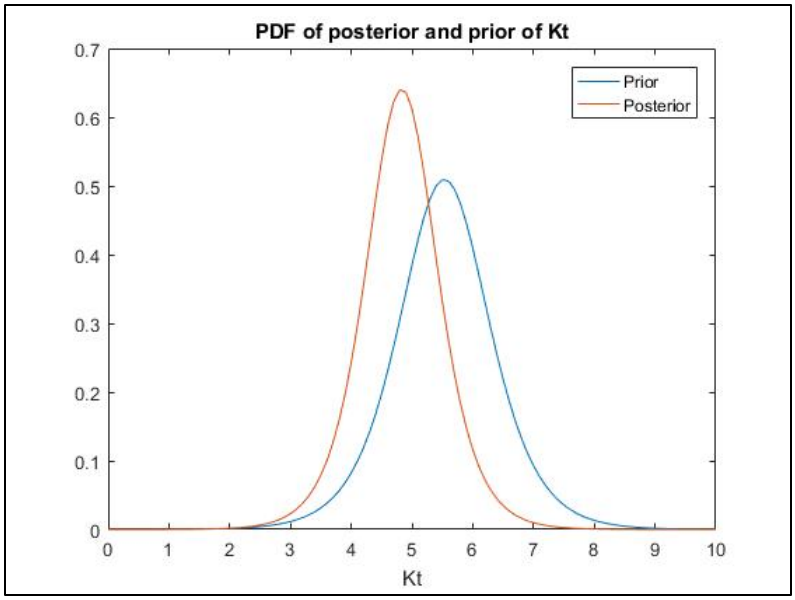


Figure 6-26. The original and updated probability of Kt

The updated distribution of Kt is then used into the MC method to generate an updated transition matrix the same way as in the previous section. The updated transition matrix is:

$$\mathbf{P}' = \begin{bmatrix} 0.6068 & 0.3922 & 0.0010 & 0 & 0 \\ 0 & 0.5072 & 0.4726 & 0.0184 & 0.0017 \\ 0 & 0 & 0.3764 & 0.5882 & 0.0354 \\ 0 & 0 & 0 & 0.4098 & 0.5902 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \textbf{Equation 6-43}$$

We assume that the observation was made at the end of the 4th time step and the optimized results are shown in Figure 6-27.

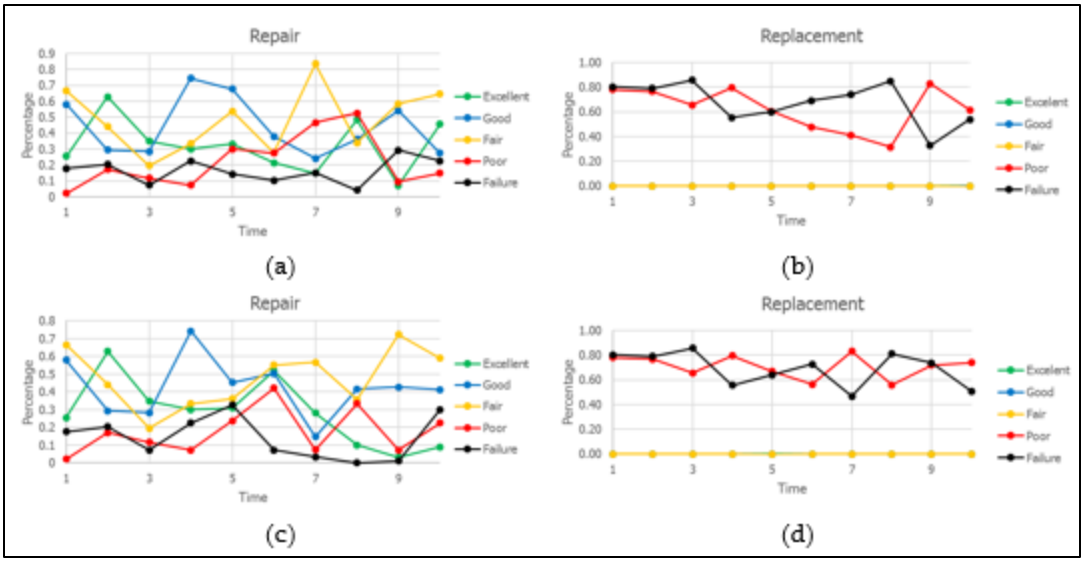


Figure 6-27. The maintenance plan with updating at the end of the 4th time step (a) and (b), comparing with the plan without updating (c) and (d)

Figure 6-28 shows the comparison of cost for maintenance with and without updating. As we can see, the cost after updating is reduced. This is due to the reduced uncertainty via the Bayesian updating.

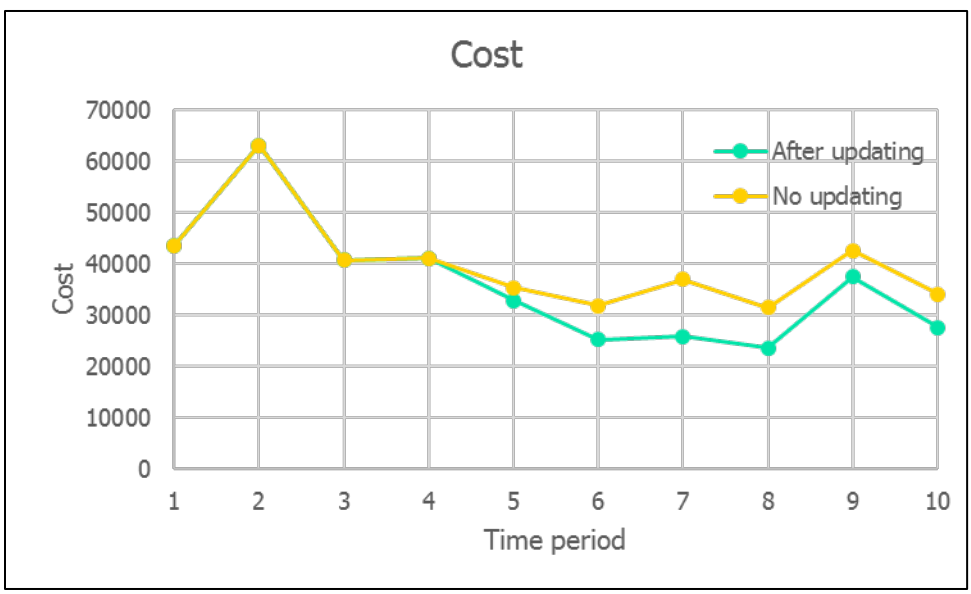


Figure 6-28. Comparison of cost for maintenance for each period

Assume that over the 10 maintenance periods, observation was made to continuously update the model parameter *Kt*. Figure 6-29 shows the change in the distribution of *Kt* after each update. As shown in the graph, the variance of the distribution decreased after each

update. After each update, the probability transition matrix can be recalculated for the next iteration.

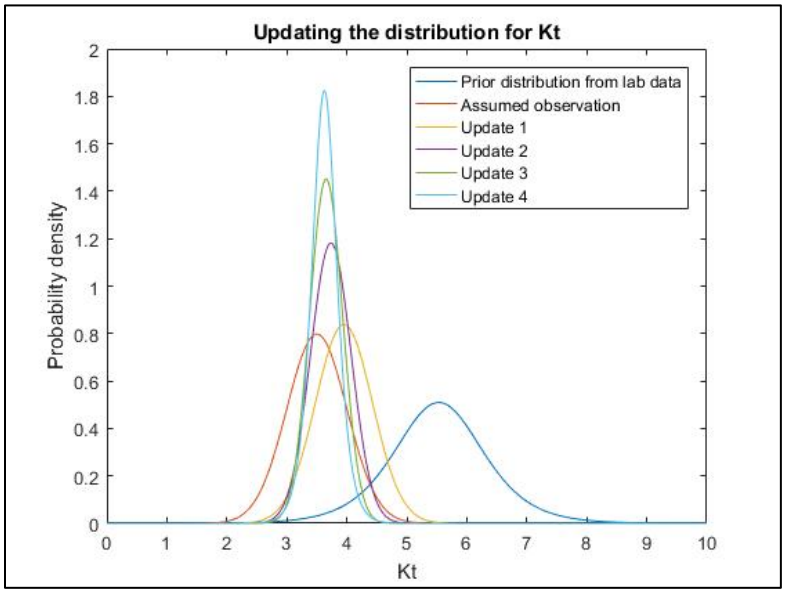


Figure 6-29. Continuous updating for parameter Kt

Following the previous example, the effect of continuous updates can be studied. Assume that the updating was done every 2 maintenance periods. Figure 6-30 shows the optimized cost for each period. According to the simulation, it is showing a trend in decrease of cost as more updating were done. The total cost for the 10 maintenance period as the number of updates was shown in Figure 6-31. The curve showed a converging trend was the number of updates increases. This is true because the variance of the parameter Kt is also converging to a very small number according to the Bayesian updating scheme.

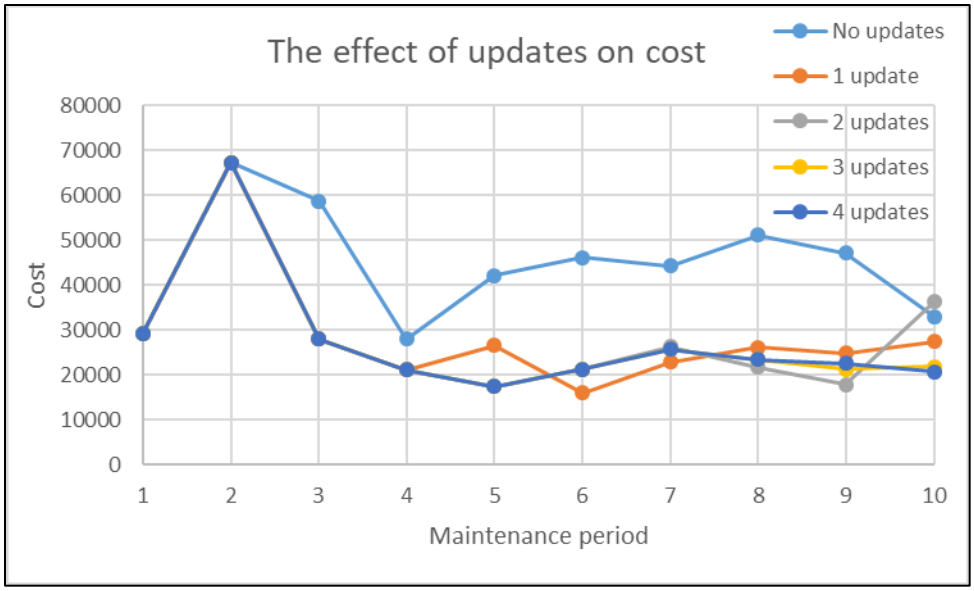


Figure 6-30. The effect of continuous updates on periodical maintenance cost

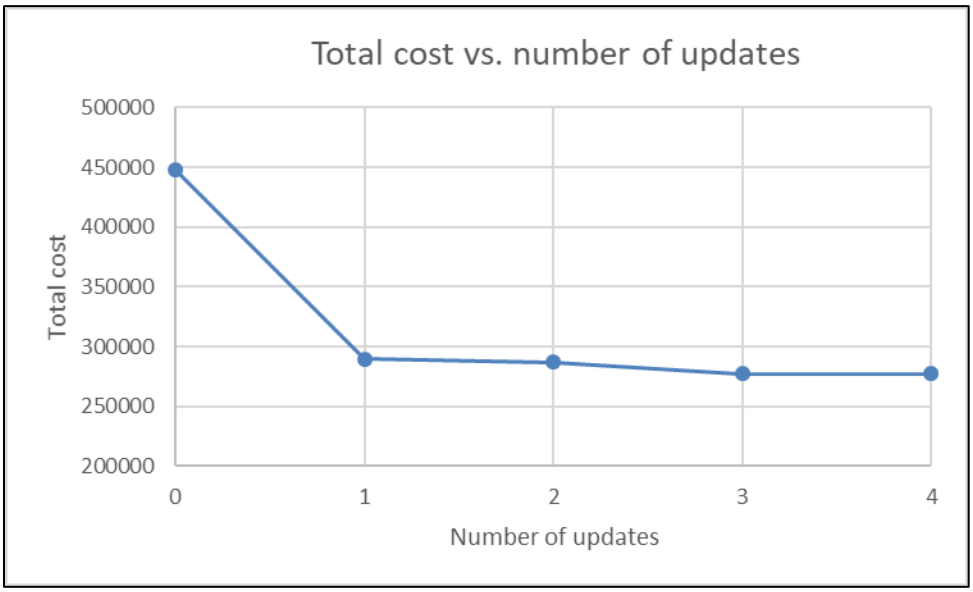


Figure 6-31. Total cost for maintenance in 10 period vs the number of updates.

The Bayesian updating framework fused the diagnostic result from the classification with the prognostic model. Dynamic reliability-based maintenance optimization can be achieved with the proposed maintenance framework.

Maintenance framework conclusion

A condition based maintenance framework were formulated for the maintenance planning of the gas pipeline system. The proposed method categorizes the pipes into different condition stages according to the crack length. The iteration of the condition vector was calculated using the degradation matrix. Given a load and a maintenance period, the

degradation matrix can be calculated from the crack growth curve. The optimization was done by limiting the value of the last element in the new condition vector, which indicates the failure probability. The fusion of prognosis with diagnosis is achieved by the real-time Bayesian updating of the model parameters with diagnostic results. The effect of the Bayesian updating reduced the uncertainty in the model parameter. The maintenance framework can also consider the consequence cost according to the location of the system.

Conclusions and future work

Accurate damage detection and classification as well as the remaining useful life prediction of infrastructure such as gas pipelines is important in maintaining the system functionality. A good maintenance plan can reduce the failure probability and ensure the infrastructure is always in working condition.

The study proposed a non-destructive damage detection method based on optical image. A device with an endoscope camera and a structured light projector was used to recode video/photo frames along the pipe. The inner surface of the pipe can be reconstructed in 3D via the video. Then, the defects in the pipe can be isolated and geometric features can be calculated. These features were put into a Naïve Bayes classifier for classification. The accuracy can be as high as 90% for the simulated data, which is comparable to more advanced machine learning algorithms such as Neural Network. The actual test was performed with the prototype device from cooperating institute. Due to the disadvantage of the prototype device and algorithm, the classification result was around 77%.

A novel BEN as a classifier was proposed to enhance the classification accuracy. It combines the maximum entropy method with Bayes' theorem to encode additional information into the network. The extra knowledge about a feature would enable a fast learning for the network. The presented work showed detailed derivation for extra knowledge given in the form of moment constraint. A special case for a Gaussian node with first order moment constraint were analytically derived. A simple example showed that the BEN classifier performs better than the NB when the training size is small and the accuracies converged as the training size increased. This is because that when large number of data is available, the trained network will be closer to the ground truth, hence the effect of the constraint may become negligible.

A simple model for the creep crack growth prediction in polymer materials were developed. The model used a power law equation and considered the damage in pipes as stress risers. The model was calibrated and validated with the lab data provided by GTI. With some proper assumptions, the model prediction agrees with the data well. Due to the stochastic nature of the crack growth process, uncertainties were introduced as the random distribution for the model parameters. The RUL of the material can be analyzed in a

probabilistic sense. Monte Carlo method is used for the calculation of failure probability given a stress level and current stage.

A condition based maintenance framework were formulated for the maintenance planning of the gas pipeline system. The proposed method categorizes the pipes into different condition stages according to the crack length. The iteration of the condition vector was calculated using the degradation matrix. Given a load and a maintenance period, the degradation matrix can be calculated from the crack growth curve predicted. The optimization was done via genetic algorithm by limiting the value of the last element in the new condition vector, which indicates the failure probability, and minimizing the overall cost. A weighing factor were added to consider the consequence of failure in different location. The result showed that more money will be spent in maintaining the functionality in areas with high consequence costs.

The fusion of prognosis with diagnosis is achieved by the real-time Bayesian updating of the model parameters with diagnostic results. According to the image analysis, the damage type and the dimension measurement of the damage can be possible. This information is used to update the parameters' distributions in the creep crack growth prediction model. The updating process can reduce the uncertainty in the model, hence change the degradation matrix in the maintenance framework. A demonstration example showed that the updating process can decrease the uncertainty in the model, hence increase the prediction accuracy of the crack growth behavior. This could help the maintenance planning to reduce the unnecessary costs or avoid unwanted failure. The information fusion between diagnostics and prognostics can achieve a more accurate risk assessment and maintenance planning.

Future work

The presented work meets the goal in the proposal statement. Some additional work can be done to enhance and further develop the results in this study:

1) The algorithm for the image reconstruction can be adjusted. The camera needs to be calibrated so that it can calculate the coordinate of the patterned light relative to the camera. With a calibrated camera, the size of the damage can also be known which is useful in the proposed dynamic maintenance framework. The dimensional information could be used to update the stress concentration factor and hence update the maintenance plan.
2) Although the proposed creep prediction model agrees well with the experimental data, there still exists some bias from the lab data. Further demonstration is needed if there were additional data available from GTI.
   The ductile group in the data set was not used, but a trend of smooth transition can be observed in the data point plot (Figure 6-15) from ductile data to SCG data. An

asymptotic function regarding yield strength of the material can be modeled for the slope of the regression curve. This would need further research and justification.

3) The maintenance framework is flexible in solving the maintenance planning given different conditions (minimize cost given reliability constraint and optimize reliability given budget constraint). The computational cost would be enormous when applied in large scale systems with various groups of pipes. The underlying principle when dealing with large scale optimization is to be studied. The maintenance framework may need to be adjusted for the high dimensional problem.
4) The proposed Bayesian entropy network (BEN) as a classifier can achieve fast learning with the extra information. The BEN concept can also be used in updating probabilities. The first step is applying the network in the updating of the parameters in the creep crack prediction model. The BEN can also be applied into large scale systems. In a large network, the update of one parameter would affect the nodes in the whole system. Sometimes this effect is not wanted. With the additional constraint, the BEN updating can be more accurate and efficient than traditional Bayesian network updating.

   Sometimes, there could be a bias between the expert opinion (empirical information) and the ground truth. Once the wrong information were coded into the network, instead of increasing the inference accuracy, it would degrade the performance of the network and even damage the whole system. Since the constraint in a BEN framework is strong, an adaptive BEN is needed to compensate this situation. The adaptive network would use the given constraint when there is not enough data to update the belief. But would change to the truth from data when more observation become available. This could be understood as when there is limited information, the network chooses to believe experience, but shifts its belief to the truth brought by data when more evidence become available.
5) The current framework focuses on one failure mode of plastic pipelines (slow crack growth). Information fusion and big data analytics with multiple failure modes and large systems that can integrate with the ongoing development of GTI framework needs further development. The “agent” serving for different failure types will need to automatically fused together for a consistent risk assessment. Recent advancement in artificial intelligence (AI) such as deep network learning has the potential for diagnostics and prognostics for gas pipeline industries.

Section 6 Bibliography

2. ISO, *9080:2012 Plastics piping and ducting systems -- Determination of the long-term hydrostatic strength of thermoplastics materials in pipe form by extrapolation*. 2012, ISO.
8. Lever, E., *Bi-Directional Shift Factors Revisited*, in *Plastic Pipes XVII*. 2014: Chicago, IL.
42. Schmalz, C., et al., *An endoscopic 3D scanner based on structured light.* Medical Image Analysis, 2012. **16**(5): p. 1063-1072.
50. O'Hagan, A. and M.G. Kendall, *Kendall's Advanced Theory of Statistics: Bayesian inference. Volume 2B*. 1994: Edward Arnold.
51. Bayes, M. and M. Price, *An essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, frs communicated by mr. price, in a letter to john canton, amfrs.* Philosophical Transactions (1683-1775), 1763: p. 370-418.
52. Jayech, K., M.A. Mahjoub, and N. Ghanmi. *Application of Bayesian networks for pattern recognition: Character recognition case*. in *Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), 2012 6th International Conference on*. 2012. IEEE.
53. Likforman-Sulem, L. and M. Sigelle, *Recognition of degraded characters using dynamic Bayesian networks.* Pattern Recognition, 2008. **41**(10): p. 3092-3103.
54. Jayech, K. and M.A. Mahjoub, *Clustering and Bayesian network for image of faces classification.* arXiv preprint arXiv:1204.1679, 2012.
55. Zhang, L. and Q. Ji, *A Bayesian network model for automatic and interactive image segmentation.* IEEE Transactions on Image Processing, 2011. **20**(9): p. 2582-2593.
56. Friedman, N., D. Geiger, and M. Goldszmidt, *Bayesian network classifiers.* Machine learning, 1997. **29**(2-3): p. 131-163.
57. Langley, P. and S. Sage. *Induction of selective Bayesian classifiers*. in *Proceedings of the Tenth international conference on Uncertainty in artificial intelligence*. 1994. Morgan Kaufmann Publishers Inc.
58. Kohavi, R. *Scaling Up the Accuracy of Naive-Bayes Classifiers: A Decision-Tree Hybrid*. in *KDD*. 1996.
59. Jaynes, E.T., *Information theory and statistical mechanics.* Physical review, 1957. **106**(4): p. 620.
60. Jaynes, E.T., *Information theory and statistical mechanics. II.* Physical review, 1957. **108**(2): p. 171.
61. Caticha, A. and A. Giffin. *Updating probabilities*. in *AIP Conference Proceedings*. 2006. AIP.
62. Landes, J. and J. Begley, *A fracture mechanics approach to creep crack growth*, in *Mechanics of crack growth*. 1976, ASTM International.
63. Williams, J. and G. Marshall. *Environmental crack and craze growth phenomena in polymers*. in *Proceedings of the royal society of London A: mathematical, physical and engineering sciences*. 1975. The Royal Society.
64. Kramer, E.J. and E.W. Hart, *Theory of slow, steady state crack growth in polymer glasses.* Polymer, 1984. **25**(11): p. 1667-1678.
65. Newman, J. and I. Raju, *Stress intensity factors for internal surface cracks in cylindrical pressure vessels.* Journal of Pressure Vessel Technology, 1980. **102**(4): p. 342-346.

66. Liu, Y. and S. Mahadevan, *Fatigue limit prediction of notched components using short crack growth theory and an asymptotic interpolation method.* Engineering Fracture Mechanics, 2009. **76**(15): p. 2317-2331.
67. Jardine, A.K., D. Lin, and D. Banjevic, *A review on machinery diagnostics and prognostics implementing condition-based maintenance.* Mechanical systems and signal processing, 2006. **20**(7): p. 1483-1510.
68. Ma, J. and J.C. Li, *Detection of localised defects in rolling element bearings via composite hypothesis test.* Mechanical Systems and Signal Processing, 1995. **9**(1): p. 63-75.
69. Davies, C. and R. Greenough. *The use of information systems in fault diagnosis*. in *ADVANCES IN MANUFACTURING TECHNOLOGY-CONFERENCE-*. 2000. TAYLOR & FRANCIS LTD.
70. Siddique, A., G. Yadava, and B. Singh. *Applications of artificial intelligence techniques for induction machine stator fault diagnosis*. in *Proceedings of 4 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics, and Drives*. 2003.
71. Loparo, K.A., et al., *Fault detection and diagnosis of rotating machinery.* IEEE Transactions on Industrial Electronics, 2000. **47**(5): p. 1005-1014.
72. Howard, I., S. Jia, and J. Wang, *The dynamic modelling of a spur gear in mesh including friction and a crack.* Mechanical systems and signal processing, 2001. **15**(5): p. 831-853.
73. Banjevic, D. and A. Jardine, *Calculation of reliability function and remaining useful life for a Markov failure time process.* IMA journal of management mathematics, 2006. **17**(2): p. 115-130.
74. Vlok, P.-J., M. Wnek, and M. Zygmunt, *Utilising statistical residual life estimates of bearings to quantify the influence of preventive maintenance actions.* Mechanical systems and signal processing, 2004. **18**(4): p. 833-847.
75. Kwan, C., et al. *A novel approach to fault diagnostics and prognostics*. in *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on*. 2003. IEEE.
76. Yam, R., et al., *Intelligent predictive decision support system for condition-based maintenance.* The International Journal of Advanced Manufacturing Technology, 2001. **17**(5): p. 383-391.
77. Guan, X., R. Jha, and Y. Liu, *Model selection, updating, and averaging for probabilistic fatigue damage prognosis.* Structural Safety, 2011. **33**(3): p. 242-249.
78. Wang, W.-b., *A model to determine the optimal critical level and the monitoring intervals in condition-based maintenance.* International Journal of Production Research, 2000. **38**(6): p. 1425-1436.
79. Peng, T., et al., *Integrated fatigue damage diagnosis and prognosis under uncertainties*. 2012, NATIONAL AERONAUTICS AND SPACE ADMINISTRATION MOFFETT FIELD CA AMES RESEARCH CENTER.
80. Peng, T., et al., *Probabilistic fatigue damage prognosis of lap joint using Bayesian updating.* Journal of Intelligent Material Systems and Structures, 2015. **26**(8): p. 965-979.
81. Gould, R.G., et al. *Investigation of a video frame averaging digital subtraction fluoroscopic system*. in *Proc. SPIE*. 1981.
82. Elgammal, A., D. Harwood, and L. Davis, *Non-parametric model for background subtraction.* Computer Vision—ECCV 2000, 2000: p. 751-767.

83. Demuth, H. and M. Beale, *Neural Network Toolbox For Use with Matlab--User'S Guide Verion 3.0.* 1993.
84. Giffin, A. and A. Caticha. *Updating probabilities with data and moments*. in *AIP Conference Proceedings*. 2007. AIP.
85. Guan, X., R. Jha, and Y. Liu, *Probabilistic fatigue damage prognosis using maximum entropy approach.* Journal of intelligent manufacturing, 2012. **23**(2): p. 163-171.
86. Skilling, J., *The axioms of maximum entropy.* Maximum-Entropy and Bayesian Methods in Science and Engineering, 1988. **1**: p. 173-187.
87. Haldar, A. and S. Mahadevan, *Probability, reliability, and statistical methods in engineering design*. 2000: John Wiley.
88. Mitchell, M., *An Introduction to Genetic Algorithms*. 1998: Bradford Books.

7. Enterprise Decision Support System (EDSS) Framework

Natural gas distribution companies with significant mileage of vintage plastic piping (specifically Aldyl A) in their distribution systems are faced with challenges in identifying segments of pipe that are at high risk of failure. This is due to several interacting factors that increase the likelihood of Slow Crack Growth (SCG) occurring. It is essential to correctly assess the fitness for service of the system as a whole, and locally where risk factors may be higher than the system's average. A holistic approach to identifying interacting factors through pipeline inspection and the development of risk models to integrate the effects of interactions and provide meaningful input into fitness for service determinations and potential mitigative strategies is needed.

In this project, an integrated set of quantitative tools were developed to support enterprise decision making process to reducing operational risk in vintage plastic distribution systems susceptible to SCG failure. The toolset allows operators, regulators and utilities to apply science and engineering methods on a variety of data sources related to pipeline distribution for fitness of service evaluation, calculate threat severity levels, and continuously monitor threat interactions and flag concerns at trigger points. The data sources may contain available system information including external conditions, inspection and leak records, historic data, customer data, and subject matter expertise.

The enterprise decision support system consists of risk models developed with Bayesian network and semantic ontology, and smart forms for field data collection. Ontology describes domain knowledge and is designed by subject matter expert to model the threat interactions. Bayesian network is ideally suited for evaluating interacting threats, investigating root causes, and predicting the effect of mitigation strategies based on conditional probabilities calculated from available data. It calculates of the probability of failures of plastic pipes and classifies them into various risk levels. The risk assessment also includes a probabilistic estimate of the remaining effective lifetime of individual segments of vintage plastic pipe. Smart forms are mobile based data collection forms that ask operators "smart" questions needed to gather relevant information for improved threat identification and risk assessment process. Specifically, a smart form captures detailed and quality data required for more granular risk analysis, such as the identification of vintage plastic pipelines with higher risks.

Introduction

Many pipeline incidents are the result of multiple, interacting causes, not a single threat. Individual threats can each be at "acceptable" levels but when overlaid result in a significant threat to the pipeline or even a failure. In this project, we are developing an enterprise decision support system platform powered by Bayesian network and semantic ontology that models the interactions between factors leading to threats or failures. Ontology describes domain knowledge and is designed by subject matter expert. Bayesian network allows calculation of the probability of failures and classification of them into various categories based on risk levels. It is ideally suited for evaluating interacting threats, investigating root causes, and predicting the effect of mitigation strategies based on conditional probabilities calculated from available data. This assessment will include a probabilistic estimate of the remaining effective lifetime of individual segments of vintage plastic pipe and a yes/no determination of whether a short-term pressure test is capable of validating the maximum defect size in the system.

Probabilistic Decision Support System Design

Semantic Ontology

Semantic Ontology provides a formal, explicit specification of a shared conceptualization of a domain. It defines the concepts and entities involved, as well as the relationships between them in an application. It also facilitates knowledge sharing over heterogeneous applications. A subject matter expert may initially guide the development of ontology that provides the domain knowledge to combine the data from different sources. Ontologies serve as vocabulary in a complex environment.

In this work, we have represented subject matter expertise of pipeline threats in ontological format. **Figure 7-1** illustrates an example of such an ontology describing interactions between factors contributing to threats in vintage pipelines. Some of the factors are residual stress, wrinkle bends/miter bend, fabrication/weld quality, third party damage, external corrosion, internal corrosion, stress corrosion cracking, soil and other superimposed stresses, and unknown, hidden, incorrect operations threats.

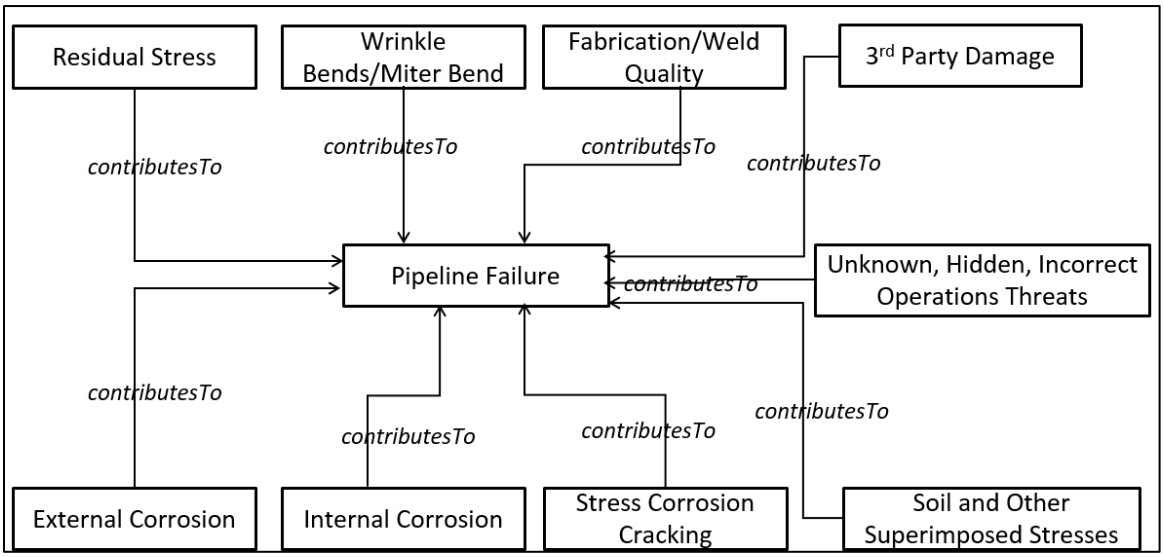


Figure 7-1. Example of ontology representing threats leading to pipeline failure

We followed an ontology design standard called *Event-Model-F* which gives a formal model of events. Event describes an action at a certain time and location. The model is based on the foundational principles of ontology and provides comprehensive support to represent time and space, objects and persons, as well as causal and correlative relationships between events. **Figure 7-2** is an ontology of failure events described by using a combination of causality and

observation patterns. It represents cracked gas pipeline event leading to the gas leakage event which is reported by an operator.

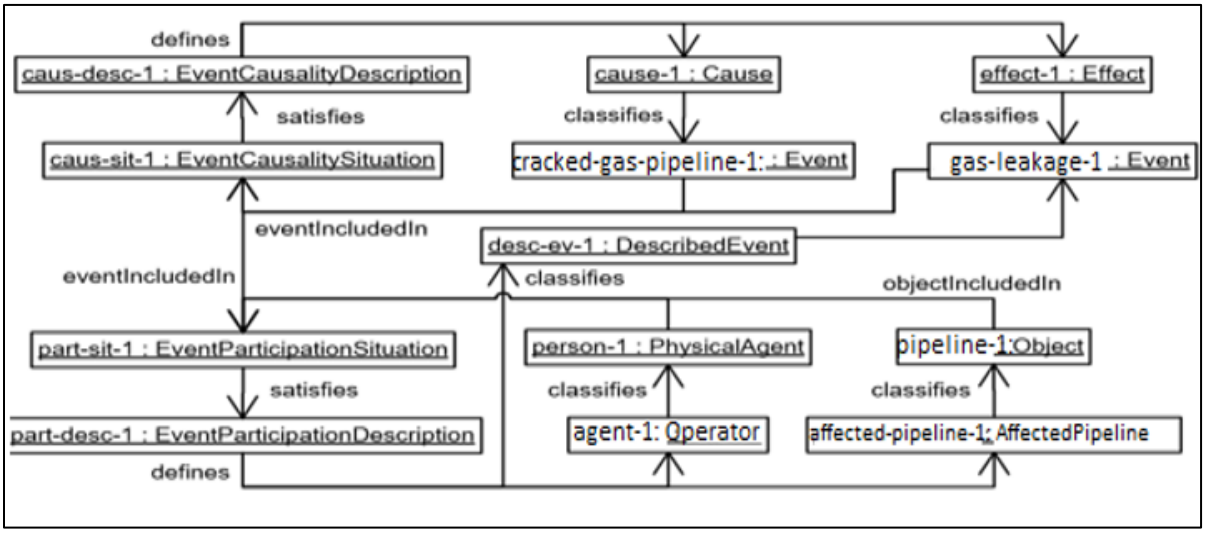


Figure 7-2. Illustration of ontology describing causality in gas distribution pipeline

Figure 7-3 is an example ontology modeling the pipeline risk computation using statistical model, subject matter expertise, historical observations, and Bayesian network.

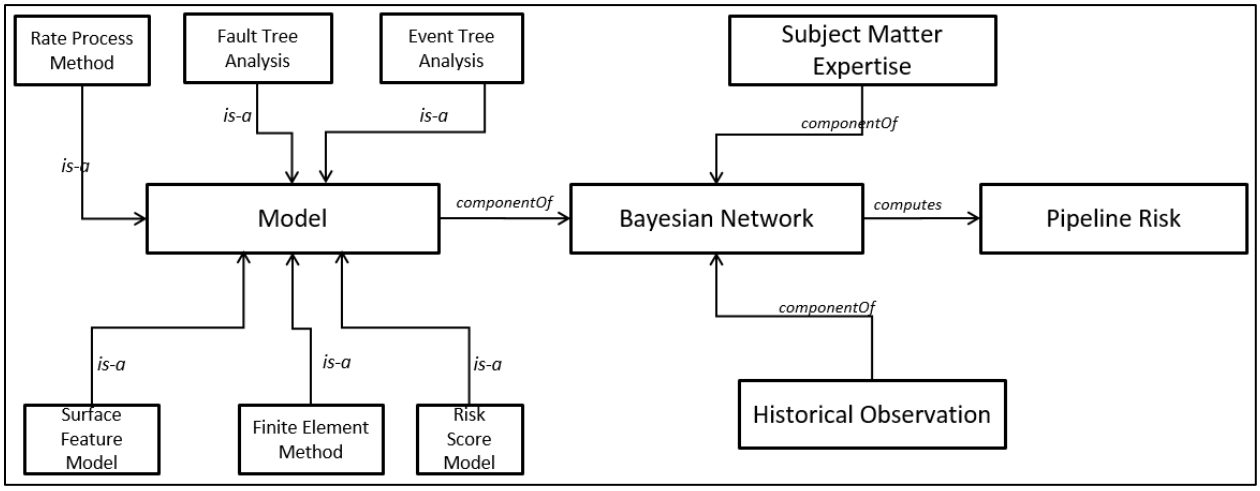


Figure 7-3. Ontology describing the models, subject matter expert's knowledge, historical observation, and Bayesian network

Bayesian foundational ontology framework was developed in collaboration with SmartCloud, Inc. to provide a probabilistic modeling approach to build domain and application knowledge. It is difficult to express incomplete, partial or uncertain knowledge in a traditional ontology. The concepts of Bayesian network model were added to the conventional ontology standard. The extended ontology follows the foundational principles of ontology and provides comprehensive support to represent time and space, objects and persons, as well as causal and

correlation relationships between events. Bayesian foundational ontology incorporates the notion of nodes, edges, states, and node probability tables as illustrated by **Figure 7-4.**

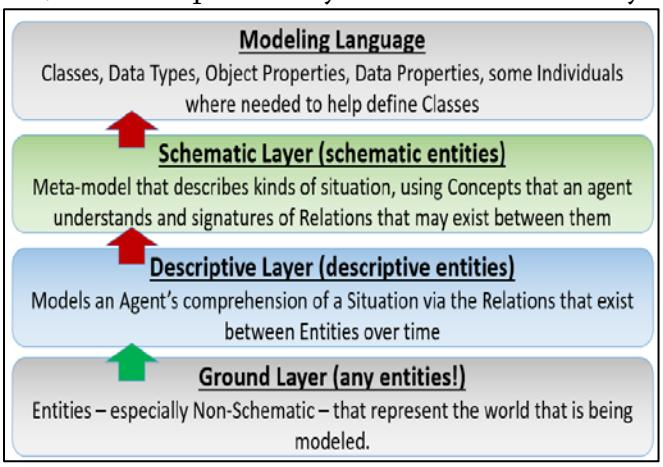


Figure 7-4. Bayesian foundational ontology

A conventional ontological approach such as in **Figure 7-5,** on the other hand, lacks the concepts needed for a Bayesian network model.

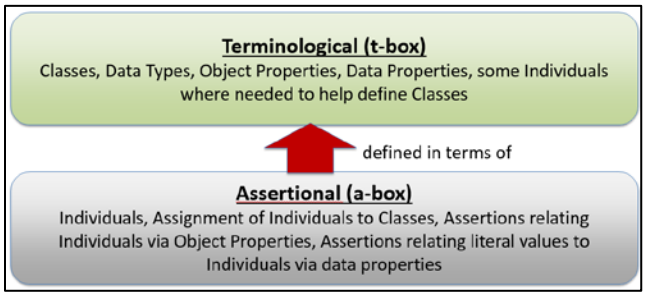


Figure 7-5. Conventional ontology

Probabilistic Enterprise Decision Support System

Probabilistic Enterprise Decision Support system (EDSS) enables operators, regulators and utilities to apply science and engineering methods on heterogeneous data sources related to pipeline distribution for fitness of service evaluation. The data sources may contain available system information including external conditions, inspection and leak records, historic data, customer data, and subject matter expertise. EDSS has models that calculate threat interaction levels and their severity. It provides a method to continuously monitor threat interactions and flag concerns at trigger points.

The conceptual design of EDSS is given below in **Figure 7-6** and **Figure 7-7**.

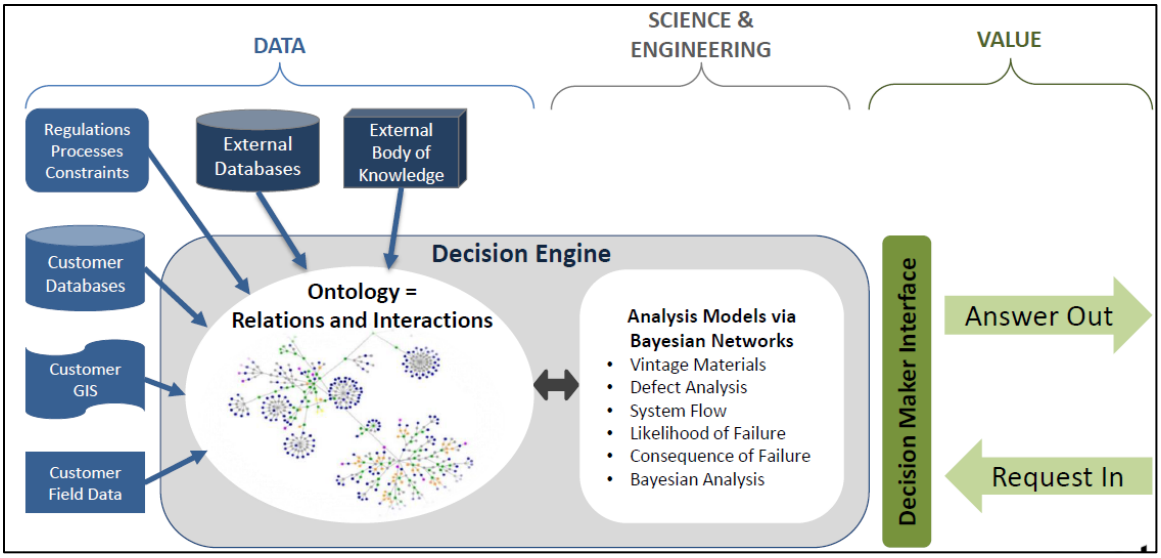


Figure 7-6. Conceptual design of Enterprise Decision Support System

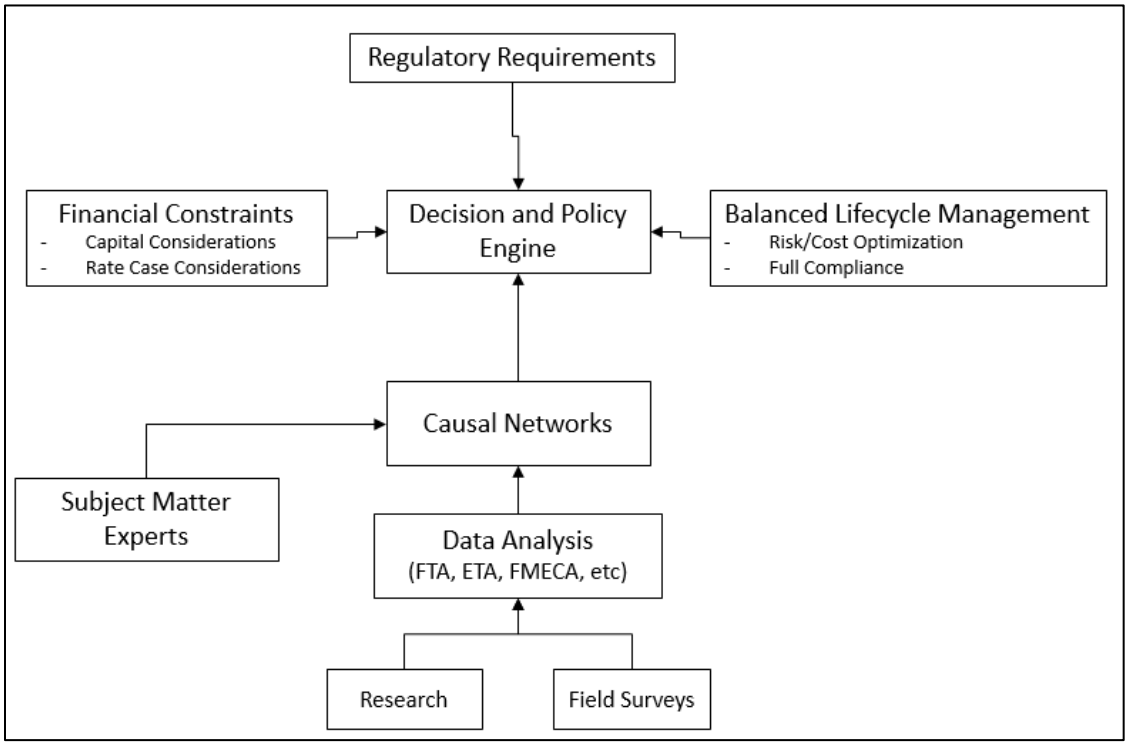


Figure 7-7. Regulations and Standards guiding decision making process in EDSS

The UML use case diagram of the EDSS is given in **Figure 7-8.** Operators, Regulators, Utilities and Subject matter experts are the stakeholders in EDSS.

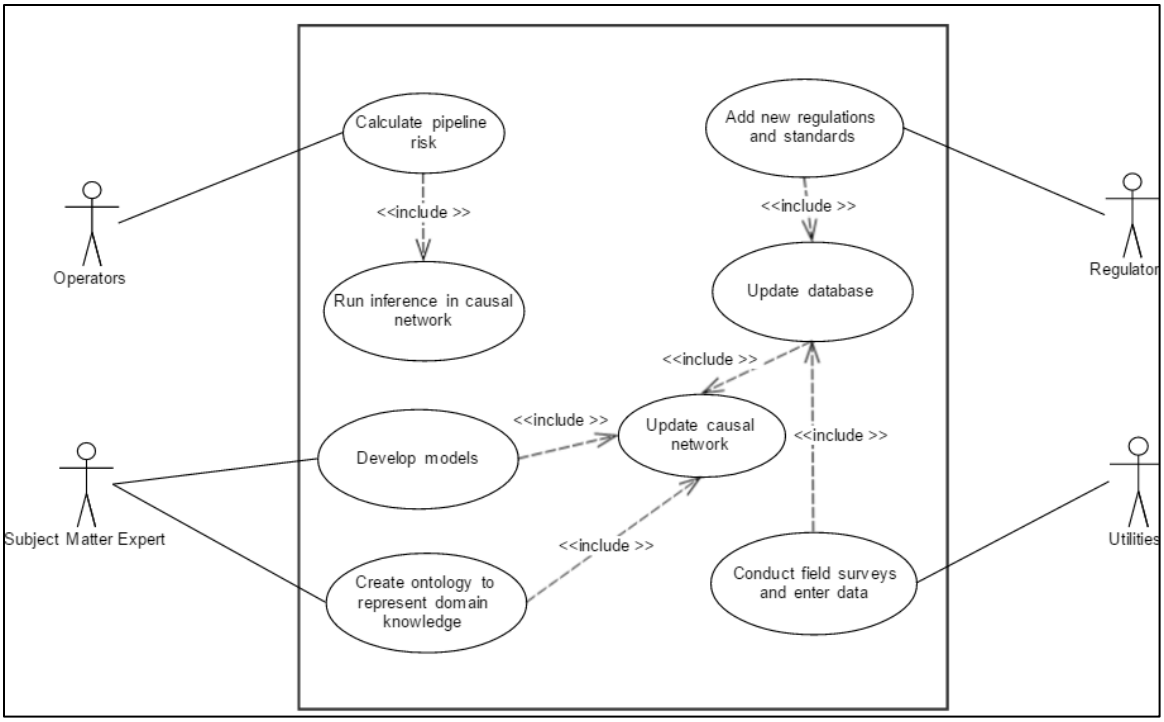


Figure 7-8. Use case diagram of EDSS

We now show the Software-as-a-Service architecture of the EDSS platform in **Figure 7-9** and **Figure 7-10**. It is a real-time, cloud-based scalable system hosted in Microsoft Azure. The heterogeneous dispersed data are stored in its big data source and semantically enabled models are deployed into the EDSS platform. EDSS is being built with open-source technologies mostly licensed under Apache 2.0 license. Two open standards have been identified for data communication in and out of EDSS: resource description framework (RDF) and JSON-LD (JavaScript Object Notation for Linked Data).

- RDF is a language for representing semantic information about objects in the World Wide Web. It is a graph based data model which provides grammar for its syntax and supports query against the model. Several design tools, including Protégé, are available for designing RDF ontology. Due to the availability of design tools and ease of usage, we have selected RDF as an ontological standard to represent subject matter experts' knowledge.
- JSON-LD is a specialized RDF syntax. As JSON-LD is also JSON document that represents an instance of an RDF data model, it fits well with the schema- less NoSQL database. Therefore, we have selected JSON-LD as communication standard between applications or models.

ISO 19465/AMQP 1.0 has been identified as an appropriate open standard messaging protocol for EDSS to efficiently transfer information within and between utilities or government agencies. It enables cross-platform applications to be built using brokers, libraries and frameworks from different vendors. AMQP 1.0 is supported in Microsoft Windows Azure and Redhat Linux operating systems. Some of its early backers are US Department of Homeland Security, Microsoft Corporation, VMware Inc, Cisco Systems, Mitre Corporation, Bank of America, Goldman Sachs, etc.

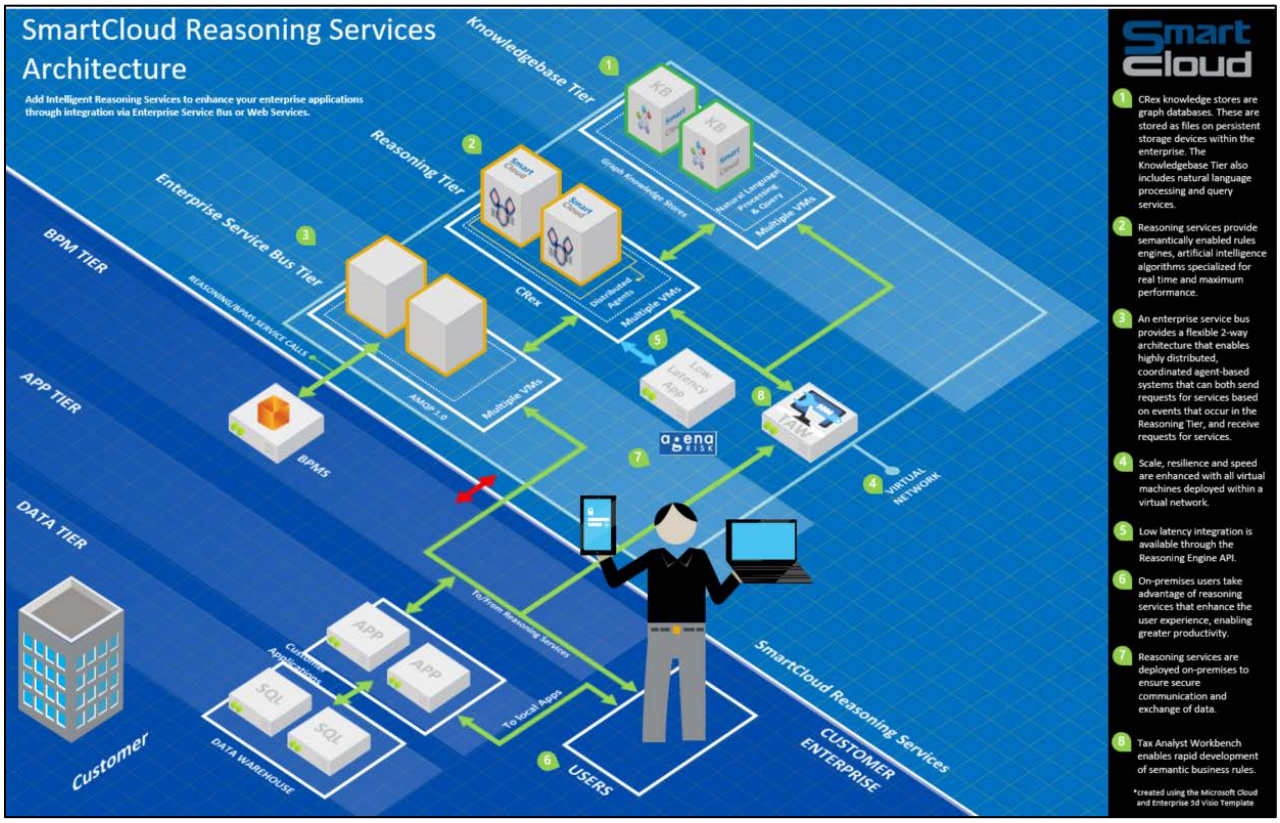


Figure 7-9. 3rd party integration of EDSS platform in their services

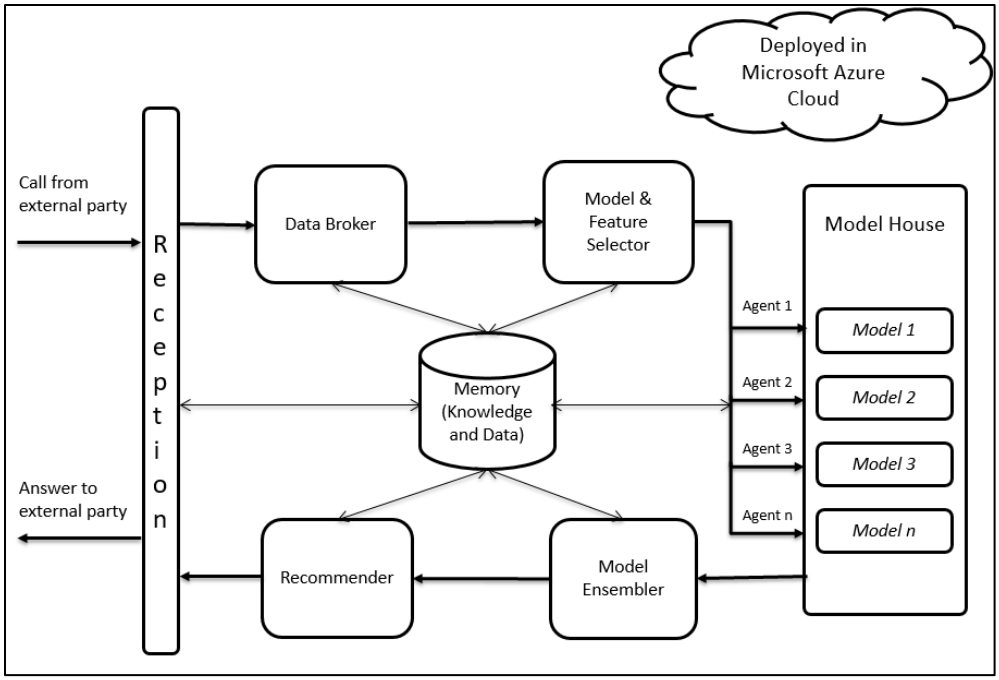


Figure 7-10. Architecture of Enterprise Decision Support System

There are nine components of EDSS platform.

Reception: It is ISO 19464/AMQP 1.0 standard communication bus which interacts with external services. It accepts queries and data as input and returns the result of causal Bayesian inference.

Data broker: It fetches information from the database. An ontological method will facilitate such data fetch.

Database/Memory: Big database technologies such as NoSQL, Graph database and Apache Hadoop are used as the database technology. NoSQL and graph database overcome the limitation of structured relational database systems by allowing schema-less data. Consequently, a variety of data in any form such as customer GIS, customer database, regulatory requirements, standards, constraints, events, sensor logs, etc are supported within the same data store. Apache Hadoop will be used as a data warehouse to ultimately store the data which may be used for analytics in future.

Model and feature selector: It performs two tasks. First, relevant single or multiple models are selected based on the query and data. Second, it identifies the important features and selects them from data by performing statistical analysis.

Model: A model is a description of the system using mathematical concepts and language. Examples of models are Bayesian network, neural network, fault tree analysis, event tree

analysis, surface feature model and finite element method, etc. Models developed with any of the programming languages such as JAVA, R, MATLAB and Python are supported.

Model House: It stores a collection of models. Graphical processing unit and central processing unit provides scalable and distributed cloud computing power to the models as required, thus supporting simple mathematical model to compute intensive image processing models.

Agent: An agent is an autonomous entity which observes events in an environment and directs its activity towards achieving goals. Each model has its own dedicated agent. The agents execute models asynchronously in parallel. In addition, it prepares data in specific format for the selected model and pulls parameter values for models from memory.

Model Ensembler: It uses multiple learning algorithms to obtain better result than could be obtained from any of the constituent models. It decides the best solution from the results of multiple models.

Recommender: It formats the final result in an appropriate format and logs that information in memory.

The design of EDSS platform as a Software-as-a-Service (SaaS) enables it to communicate with other enterprise or government applications, thereby increasing its usability and a value addition to utilities, operators, regulators and public.

Data Entry

"Smart form" was designed and developed for data entry purpose. Smart form is an electronic form with capability beyond a traditional data entry interface with a fixed schema. The fields in the form are generated dynamically based on the content and context of the data entered so far. Subject matter expertise, represented as ontology, guides the process of data entry. It will guide the operator through the relevant data gathering stages and ensure only valid and relevant data is fed directly into the appropriate models. It is designed to overcome human data entry errors, which are the most common source of bad data, and to collect as much complete data as possible.

Simulation

EDSS platform provides a set of toolsets for simulation and modeling **Figure 7-11**. It has access to heterogeneous big data sources including simulated test data, manufacturers' data, field failure data, research reports, academic literature, standards and regulations, and subject matter expertise. Operators or analysts may run simulation with models that include application specific handbook models, probabilistic risk models, and calculators. Multiple competing models may be deployed and run in parallel to evaluate the results.

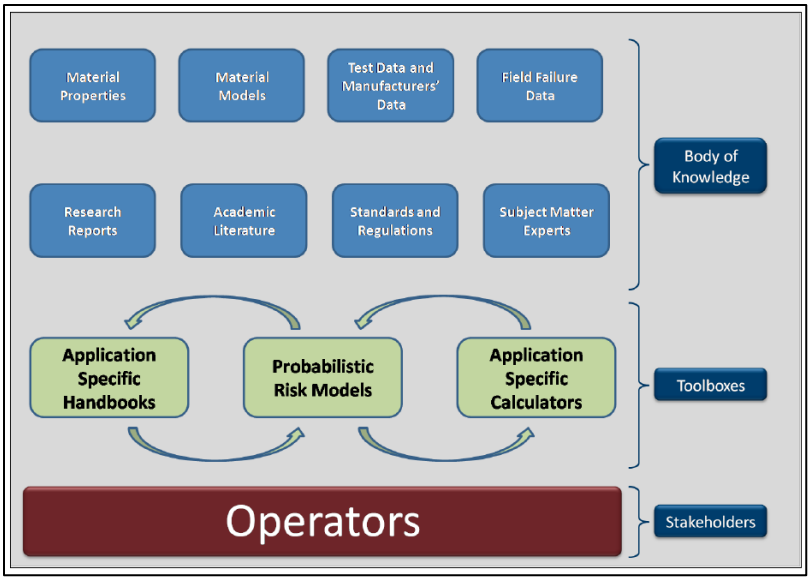


Figure 7-11. Simulation model toolboxes with access to heterogeneous data sources

Insights Interface

Insights interface **Figure 7-12** is designed as a web-based dashboard that operators or regulators can configure to gain key insights into the information they want including vital details about the data, regulations and standards. Furthermore, it allows users to select multiple risk models, enter parameters, and run them in parallel to display the comparative results as graphs and visuals.

Figure 7-12. Mockup version of insights interface dashboard

Query

A query is a request for information from the EDSS system. In this work, four different types of queries are supported – diagnostic query, predictive query, inter-causal query, and combined query. **Figure 7-13** illustrates these four types of query.

a. *Diagnostic query* performs reasoning from effects to cause. For instance, the questions such as - determine the stress concentration and temperature that reduces the pipeline life expectancy by ten years – may be answered with diagnostic query.
b. *Predictive query* performs reasoning from cause to its effects such as "what is the change in pipeline life expectancy for an increase in stress concentration by 50 psi?"
c. *Inter-causal query* involves reasoning about the mutual causes of a common effect.
d. *Combined query* provides a flexible reasoning to find intermediate results of diagnostic and predictive query.

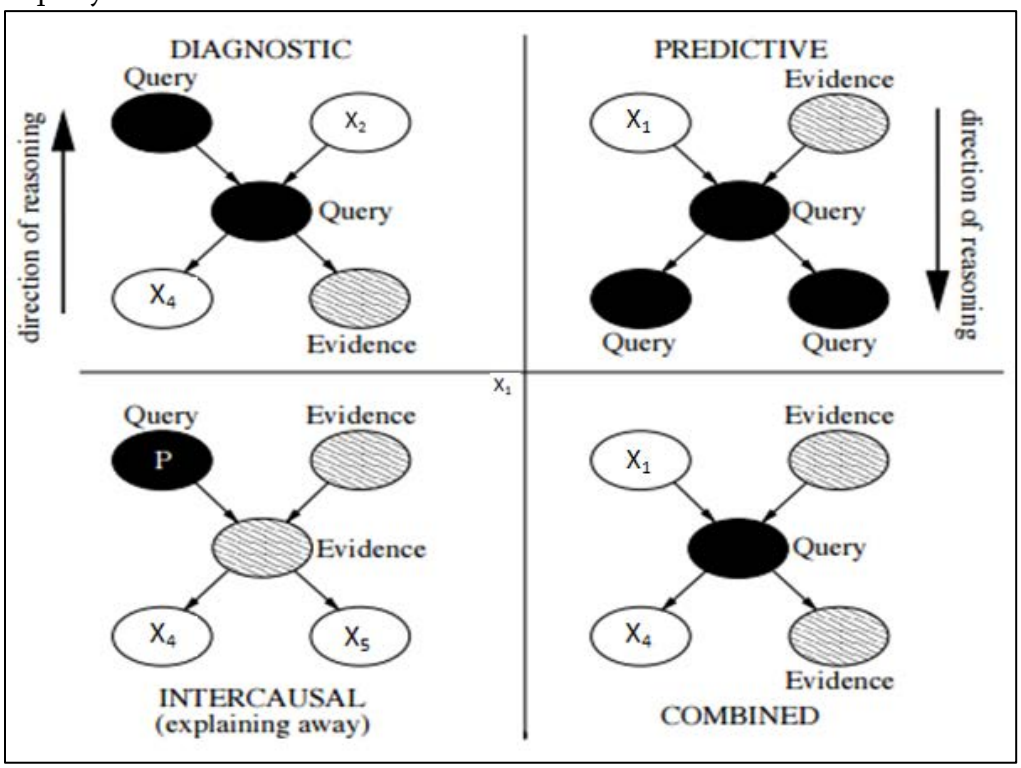


Figure 7-13. Supported Query Types

Data Communication

Figure 7-14 illustrates the data flow between user interface, database, and model modules. The data is communicated between these modules in a JavaScript Object Notation (JSON) format. JSON is a lightweight data-interchange format that is built on two structures:

- A collection of key/value pairs, similar to a dictionary entry.
- An ordered list of values

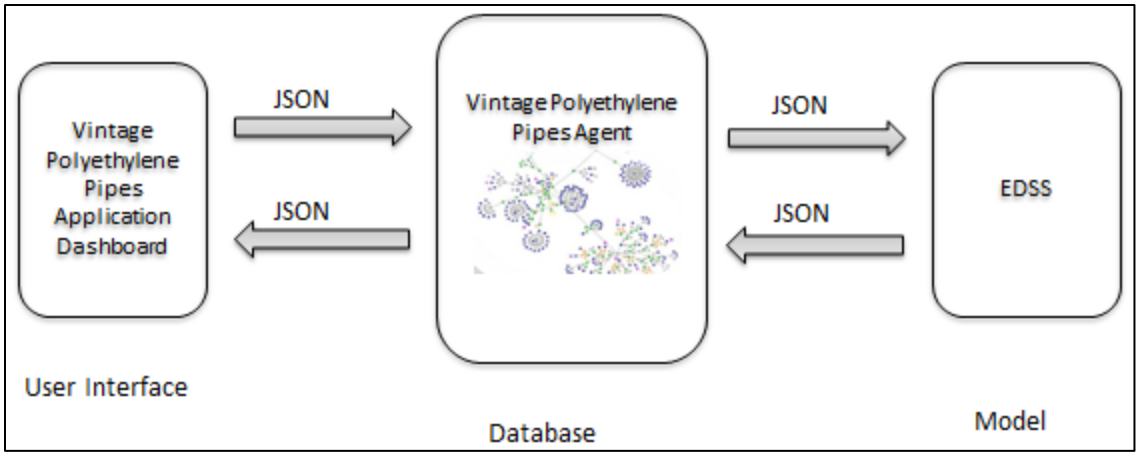


Figure 7-14. Data flow between user interface, database, and model modules in EDSS

The template of a generic JSON message developed is given below. It is used to initialize, update, and query the Bayesian network in the EDSS.

JSON message = {

```
"app": "",
"operation": "",
"bNodes":[
    //categorical ranked/unranked; Boolean; discrete node
    {
        "bName": "",
        "bType": "",
        "bNPTType":"",
        "bParents": [],
        "bNPT": {
            "bNPTValue":[
                {
                    "bNPTState":"",
                    "bNPTStateValue":0.0
                }
            ]
        },
        "bStates": []
    },
    //continuous interval node with expression representation of NPT
    {
        "bName": "",
        "bType": "",
        "bNPTType":"",
        "bParents":[],
```

```
                "bNPT":{
                    "bNPTExpressionType":"",
                    "bNPTValue":[{
                            "bNPTMean":0.0,
                            "bNPTVariance":0.0
                    }]
                },
                "bLower":0.0,
                "bUpper":0.0
            },
            //continuous interval node with partitioned expression representation of
            NPT
            {
                "bName": "",
                "bType":"",
                "bNPTType":"",
                "bParents":["","",],
                "bNPT":{
                    "bNPTExpressionType":"",
                     "bNPTValue":
                    [
                            {
                                    "bNPTParentState":          {"bNPTParent":"",
            "bNPTState":""},
                                    "bNPTArithmenticEquation":""
                            },
                            {
                                    "bNPTParentState":          {"bNPTParent":"",
            "bNPTState":""},
                                    "bNPTArithmenticEquation":""
                            }
                    ]
                },
                "bLower":0.0,
                "bUpper":0.0
            }
        ]
    }
```


The fields in the JSON message are -

1. app - Name of the application

Type: String
Required: True

2. operation – Operation type

   Type: String
   - INITIALIZE
   - QUERY
   - UPDATE

   Required: True

3. bNodes – List of nodes

   Type: JSON
   Required: True

4. bName – Name of node

   Type: String
   Required: True

5. bType – Type of node

   Type: String
   - *BOOLEAN*
   - *CONTINUOUS*
   - *DISCRETE*
   - *INTEGER-INTERVAL*
   - *CATEGORICAL-RANKED*
   - *CATEGORICAL-UNRANKED*

   Required: True

6. bParents – Parents of a node

   Type: String array
   Required: True

7. bNPTType – Type of node probability table

   Type: String
   - *NUMERIC*
   - *EXPRESSION*
   - *PARTITIONED*

Required: True

8. bNPT – Node probability table

   Type: JSON
   Required: True
   //categorical ranked/unranked; boolean; discrete node

```
"bNPT": {
        "bNPTValue":[
        {
                "bNPTState":"",
                "bNPTStateValue":0.0
            }
    ]
}
```

   //continuous interval node with expression representation of NPT

```
"bNPT":{
        "bNPTExpressionType":"",
        "bNPTValue":[{
            "bNPTMean":0.0,
            "bNPTVariance":0.0
        }]
    }
```

   //continuous interval node with partitioned expression representation of NPT

```
"bNPT":{
        "bNPTExpressionType":"",
      "bNPTValue":
    [
            {
                "bNPTParentState": {"bNPTParent":"", "bNPTState":""},
                        "bNPTArithmenticEquation":""
            },
            {
                "bNPTParentState": {"bNPTParent":"", "bNPTState":""},
                        "bNPTArithmenticEquation":""
            }
        ]
}
```

9. bStates – States of a node

   Type: String array

Required: No (except in categorical and Boolean node types)

10. bLower – Lower range of a node

    Type: Decimal
    Required: No (required in Continuous, Discrete, Integer interval node types)

11. bUpper – Upper range of a node

    Type: Decimal
    Required: No (except in Continuous, Discrete, Integer interval node types)

Probabilistic Decision Support System Development

Smart Form

Smart forms are the data acquisition forms generated automatically from ontologies, and are useful in aggregation of the semantically enriched input gathered through these forms to feed into EDSS. The smart forms developed in this project are keyhole data gathering form, first response form and audit form. The architecture of smart forms is shown in **Figure 7-15**.

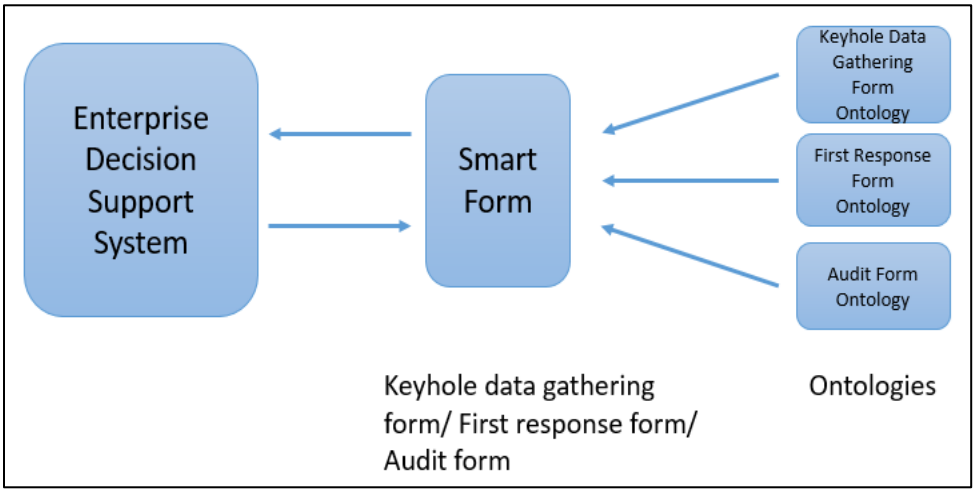


Figure 7-15. Basic Workflow Architecture of Smart Forms

Smart forms has inbuilt logic to improve the data collection process and identification of the root cause during the pipeline repair process. The logic provides "smart" questions and answer options to guide the user through the process of determining the root cause to ensure that the correct root cause is identified and appropriate supporting information is provided. The smart form logic and definitions can be integrate into utilities existing field data collection forms. A smart form improves the quality of data collected during the repair process and leads to improved threat identification and risk assessment for DIMP. Further, a smart form allows more detailed information to be captured to allow more granular analysis performed, such as the identification of key threat trends.

The use of "other" or "Not Available" during leak root cause investigations limits the ability of operators to understand trends, threats, and risk. Data collection programs such as CGA's DIRT, Plastic Pipe Data Committee, and PHMSA routinely publish reports with between 10% and 60% of root causes listed as "other" or "Not Available". The developed approach improves the data quality and ultimately the ability of the industry to identify and respond to threats in a timely manner.

Methodology

Root Cause Analysis Process

Fault Tree Analysis

FTA[8] is an ideal method for breaking the data gathering process into logical steps **Figure 7-16** and . The top gate is the fault condition and a series of and/or gates define the sequence of conditions that lead up to the fault condition. The lowest level gates can be viewed as the root cause at the desired level of granularity (the root cause may not be fully resolved).

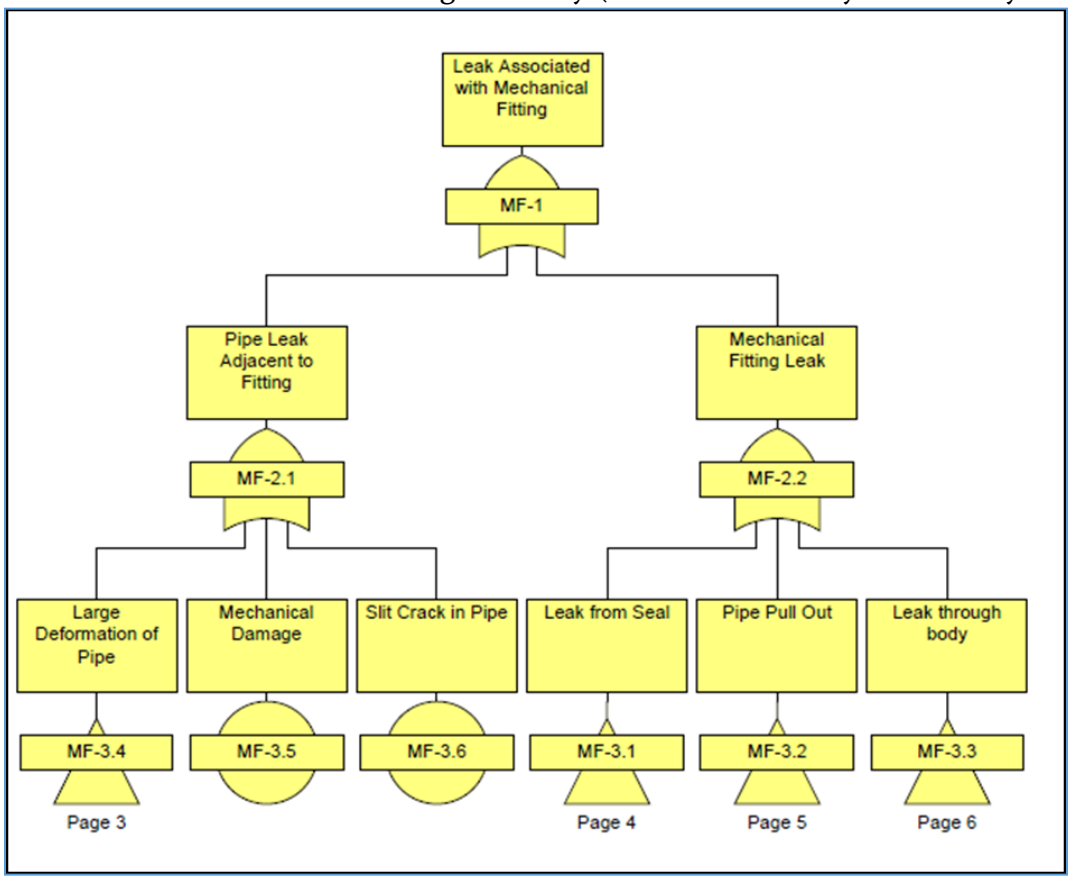


Figure 7-16. Example of a fault tree

A complete fault tree analysis, as depicted in **Figure 7-17**, can be very large.

---

8 A concise history of the FTA methodology, how it can be used and the advantages of the method is presented in Wikipedia:
"http://en.wikipedia.org/w/index.php?title=Fault_tree_analysis&oldid=628258471"

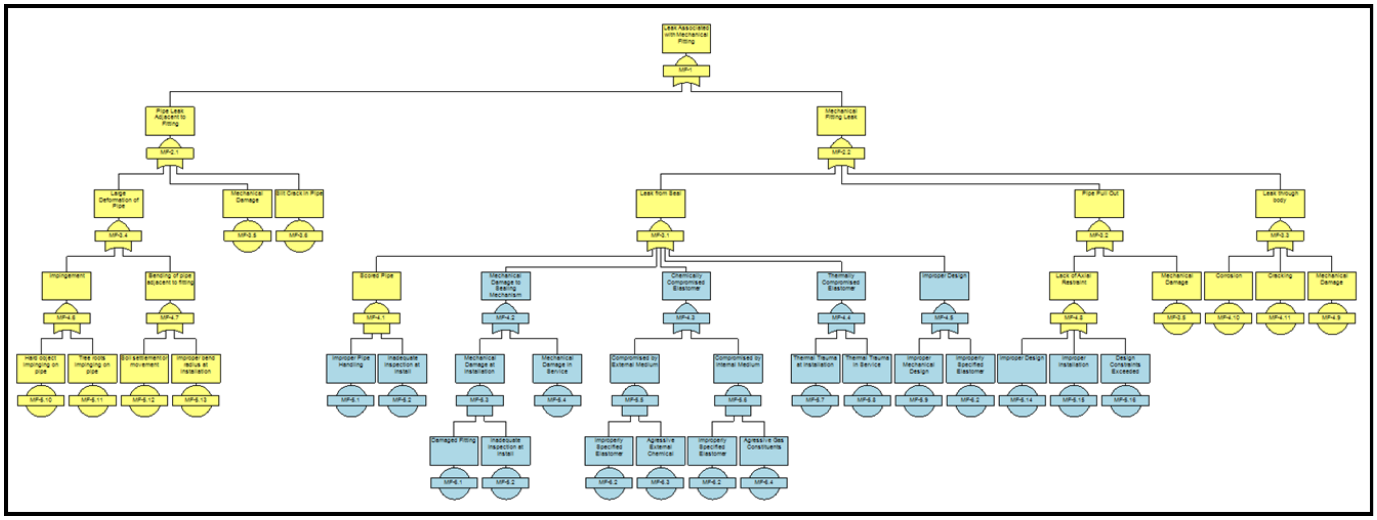


Figure 7-17. Full FTA for leaks associated with mechanical fittings

Some portions can be readily diagnosed and resolved in the field (those nodes depicted in yellow in Figure 7-17). The nodes depicted in blue will likely need a separate and more expert investigation to be properly resolved.

In order for the RCA process not to become an unwieldy, we need to come up with methods to break analysis into digestible portions. This is easily accomplished by converting the fault tree into a Decision Tree.

Decision Tree

A Decision Tree[9] is a very convenient flowchart like graphical depiction of an algorithm where the pathways from root to leaf depict classification rules **Figure 7-18** and **Figure 7-19** .

9 A concise description of Decision Tree is presented in Wikipedia: http://en.wikipedia.org/wiki/Decision_tree

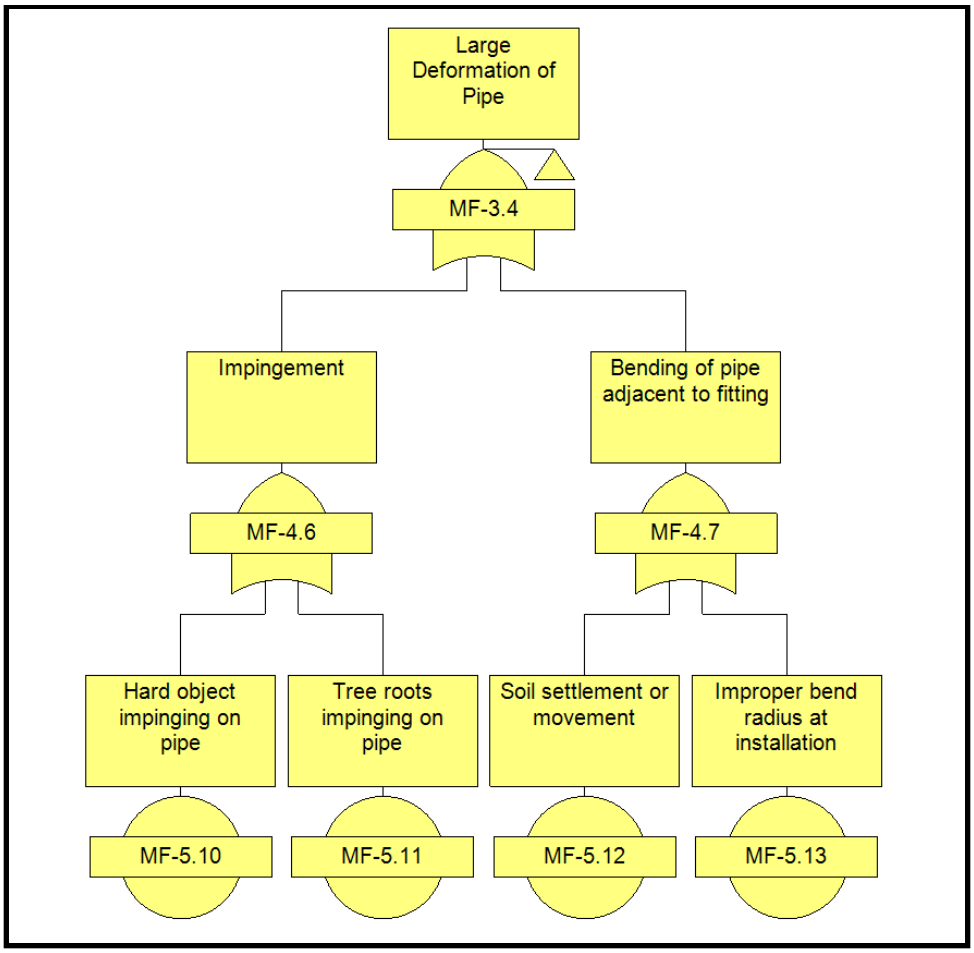


Figure 7-18. Fault Tree depiction of Decision Tree

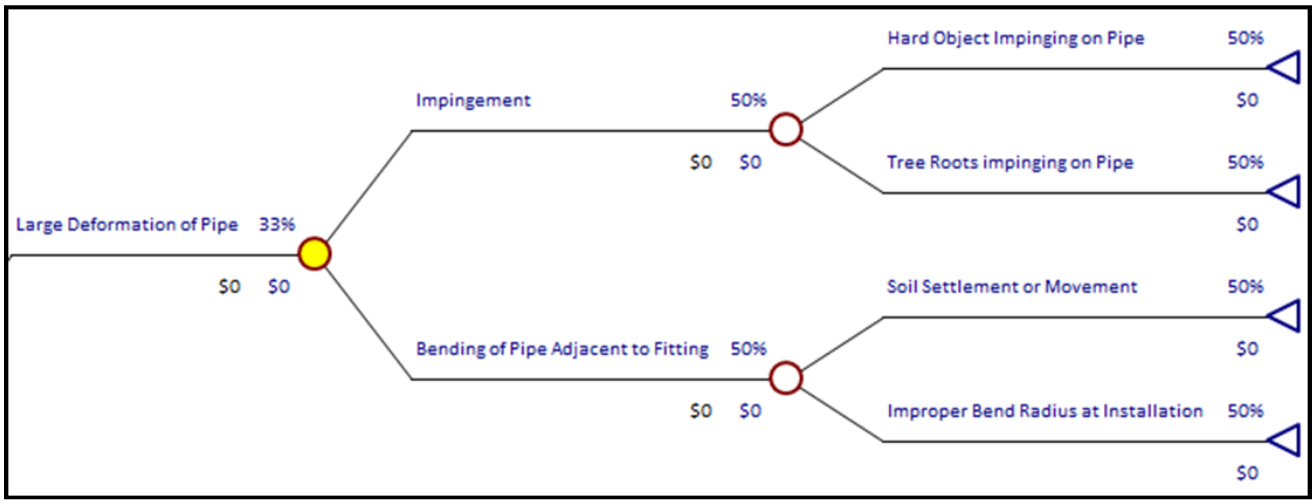


Figure 7-19. Decision Tree depiction of the FTA

A fault tree can be traversed in either direction – top level event to causes, or cause to possible top level events. The decision tree depiction is better for moving from the observed

event to the root cause through a series of binary questions – the next pathway is either activated, or it is not.

A decision tree consists of 3 types of nodes:

1. Decision nodes - commonly represented by squares
2. Chance nodes - represented by circles
3. End nodes - represented by triangles

In this application we are not making actual decisions, we are simply traversing a logical pathway, so the nodes utilized are chance nodes and end nodes. A chance node, or circle in the diagrams presented, simply indicates that there are several alternate pathways to choose from. The operator is presented with enough information to enable a quick choice as to which pathway is correct. The pathway terminates at an end node (triangle), which is the point that the subject matter experts have determined to be the best point of termination for the particular investigation for which we are gathering information.

Each branch in the **Figure 7-20** is a set of questions on the electronic form. Depending on the option that is clicked **Figure 7-21** the next screen will display the relevant choices. Relevant and useful pictorial or textual guidance can be displayed in association with each question.

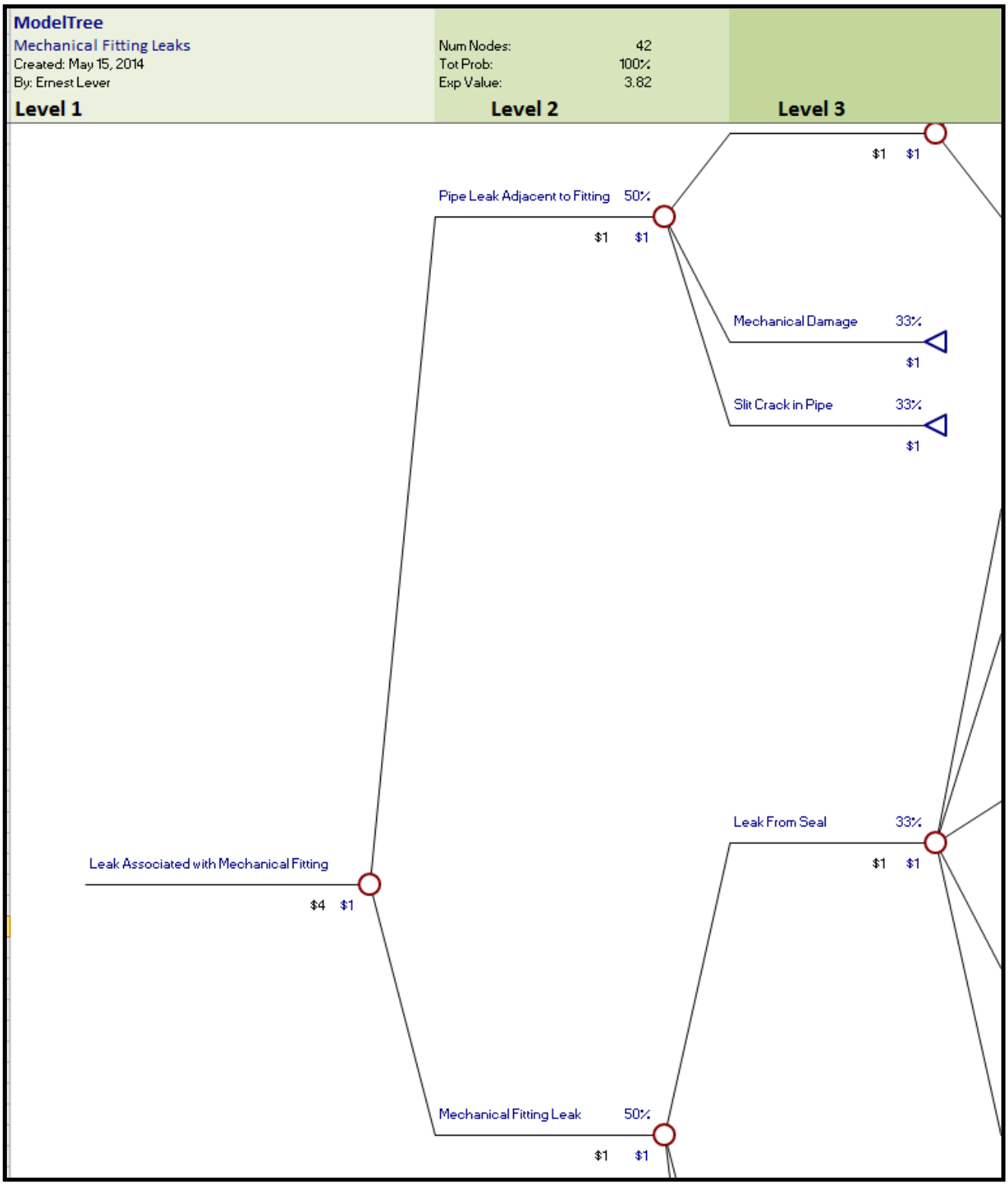


Figure 7-20. Entry level for the "Leak Associated with Mechanical Fitting" decision tree

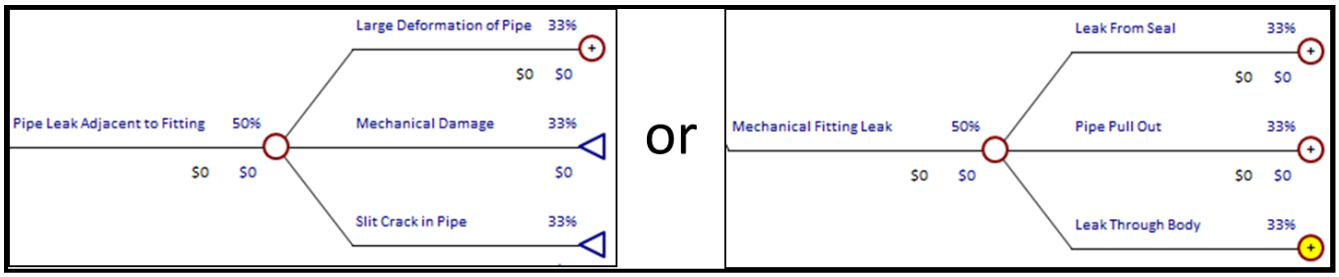


Figure 7-21. Moving along the decision tree towards the root cause

Gathering Context Sensitive Information for Risk Models

Working through the decision tree that was defined by the original fault tree we will eventually arrive at a root cause. Some nodes may require extensive investigation for the correct follow on path to be identified, but the fact that the form is derived from well-defined fault tree analyses, developed by subject matter experts, allows simply structured assistance to be provided to the operator at each node of the decision tree.

The application can prompt the operator to gather useful meta-data related to the particular investigative pathway, such as detailed photographs, videos captured at the site or any additional that is known to be helpful in a determining the implications to the rest of the system due to the incident that occurred. This information gathering prompts is driven by the collected knowledge of analyzing many similar incidents that is embedded in the fault trees and is only be provided in the proper context so as to avoid overloading the operator with reams of information that he or she has to evaluate on the fly. The prompts take the form of simple, clear instructions with pictorial assists, provided in the proper sequence to ensure efficiency of the information gathering process.

Eliminating the "Other" Category

The operator is presented with clear choices between alternate pathways at every point in the field investigation. If there is insufficient information to answer a set of questions the investigation is stopped at that point with a simple note "insufficient information available to continue". We then have a very clear record of what was available to the field crew and where they were unable to continue due to lack of resources or information. If the incident warrants further investigation the subject matter experts called in to continue the investigation have a well-defined start point for their detailed investigative work.

The application can associate a belief score for each node in a pathway that is a function of:

- How many subsequent nodes are resolved,
- How much supporting evidence for each pathway is stored in the meta-data.

Data Driven Risk and Consequence Models

The decision tree underlying the form is essentially a risk and consequence model:

- Field data for each node can be directly used to assign probability of occurrence for each pathway, and
- Operator data for associated costs can be directly input into the decision tree yielding an expected cost for all possible system faults as well as each particular branch.

This functionality will not be exposed in the smart form application, but the data gathered can be directly input into an associated risk and consequence model. Because the form structure is derived directly from a fault tree analysis the data gathered is perfectly structured to feed into risk and consequence models.

Involving Standards Organizations

An industry standard can be developed to define the accepted fault trees for all common field failure modes. The standard can define:

- The appropriate questions to be asked at each level of the process:
    - Filling out the form
    - Conducting ancillary root cause investigations
- The appropriate methods for developing supporting data
- The consensus root causes for each field failure mechanism

Benefits of Proposed Smart Form

The major benefits are:

- Well-defined, consistent and coherent logic,
- "Other" category avoided,
- Potential for industry standard,
- Readily implemented in electronic format,
- Integrate with mobile data apps,
- GIS integration possible,
- Tracking and Traceability integration possible, and
- Data collected can be fed directly into risk and consequence models (consistent structure).

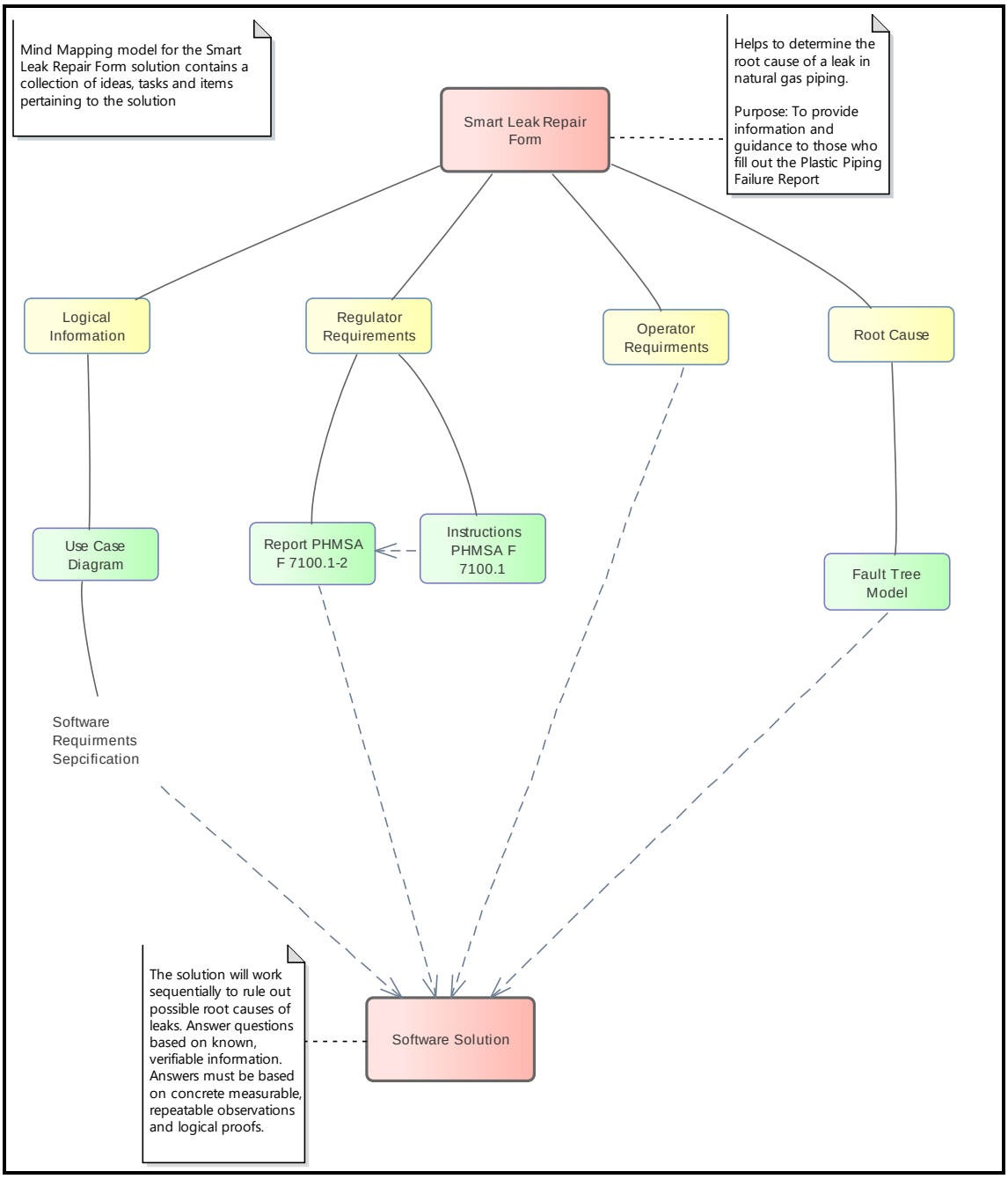


Figure 7-22. Mind Mapping Diagram of Smart Form Application

Development

The smart form application improves the data collection process by guiding operators through a series of relevant questions about the system being inspected. High-level requirements that are supported by the developed application include the following:

- Runs on an Android smart (mobile) device.
- Supports disconnected data collection in cases of no internet access.
- Supports syncing of disconnected data once internet access is restored.
- Supports flexible data entry so that any non-standard data can be captured.
- Support the auto-generation of a point location, this should represent the GNSS point where the user created the record. The point feature should be inserted into Esri Geodatabase.
- Supports taking a geotagged photo that is retained and related to the record of interest.
- Supports User Authentication including "organization" or "domain"
- Support Deployment Configurations
- Manage Data Connection
- The Smart Forms Platform enables configuration of decision-driven workflows using graphical editors that map specific workflows to forms. The architecture for this platform is shown in **Figure 7-23** Architectural and software selection is based on flexibility and agility through maximizing application embedded intelligence and creating interfaces that are intuitive to field users.

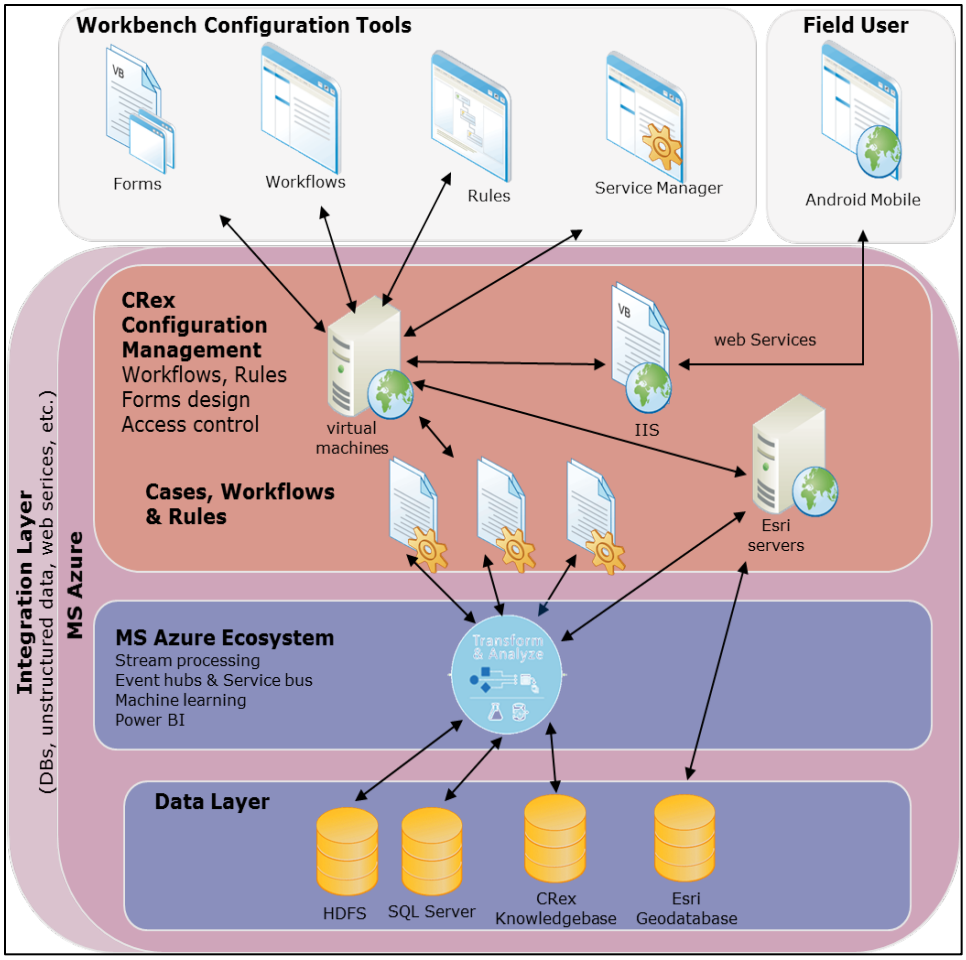


Figure 7-23. High level systems architecture developed for the Smart Forms Platform

The implementation is a modern, robust and fully-configurable Smart Forms creation platform for Android mobile devices. The platform is provided as a cloud-based service with developer, configuration and end-user (field user) access to different features. Central to the platform is a Configuration Manager application that enables use case configuration for new forms-driven applications through graphical editing of displayed forms as well as the workflows that drive the deployed smart forms. The platform takes advantage of the ability of the platform to map individual forms to tasks or activities within rules-driven workflows. A case management feature, essentially a workflow that manages other workflows, is provided to manually or automatically select from among a number of workflows to implement in the mobile application and that drive display of forms. Workflows are constructed visually within a Workbench and specific tasks within a workflow direct the display of forms in an intelligent way using rules entered through structured natural language editors, also within the Workbench. Other methods can be employed to make the decisions required by smart forms. For example, a Bayesian network might be evaluated in a task within a workflow and the results used to decide which additional forms to display or the data to display within a form. The capabilities are generalized in such a way that it is not necessary to upgrade, recompile or

redeploy the Smart Form application in order to release new use case configurations into production. Only the configuration needs to be changed, and if new intelligent services are provided in the future, these can be accessed through standard connectors. A Service Manager is provided for user authentication based on organization or domain and configuration management **Figure 7-24**.

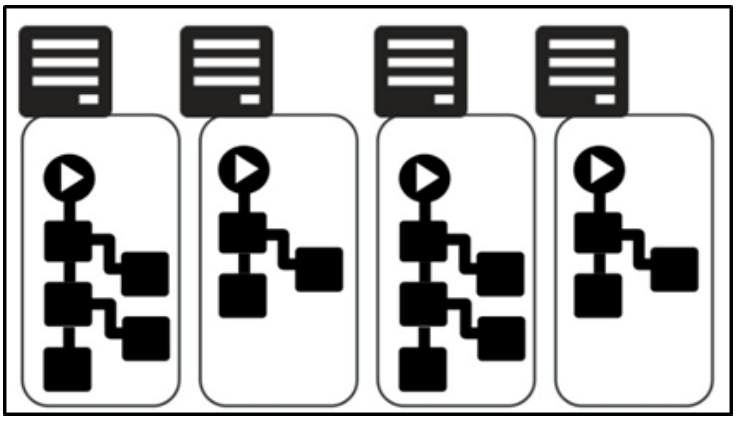


Figure 7-24. Configuration of forms to specific workflows

Development and deployment within Microsoft Azure allows access to other services within the Azure Transform & Analyze ecosystem, including Stream processing, Event Hubs, Service Bus, machine learning, Hadoop and Power BI. Of these services Service Bus and Power BI is used in this project. Future efforts can take advantage of all Microsoft services, and others created by Microsoft or 3rd-party vendors in the future. The semantically-enabled Bayesian network engine developed with GTI in an effort parallel to this project, is provided within this layer. The architecture is designed to meet the immediate Smart Forms project needs, while allowing for future expansion and eliminating obsolescence.

The data layer includes the CRex semantic knowledgebase and SQL Server for raw or unstructured data storage. CRex is also able to use SQL Server to store semantically-represented knowledge, typically historical information for archival. An Esri geodatabase in this layer provides for storage of geospatial information, obtained from the Android smart device, or other external services through the integration layer.

An integration layer is provided for programmatic interface to the Smart Forms Platform. There is one primary connection technology to support for this integration; AMQP 1.0, an enterprise service bus standard with connectors available to a large number of structured and unstructured data sources and targets.

Components

SmartForms improves both the collection and analysis of pipeline network integrity data through the development of a proof of concept (POC) server/mobile device application. The SmartForms POC consists of two applications, a locally installed mobile device app for field use and a host server application that stores knowledge, enable analysis, and intelligently drive field data collection.

The mobile device application will be used by field personnel. The mobile app uses a series of workflow-driven forms to guide field technicians through the collection of data. The mobile app supports both connected and off-line operation. When connected, data collection is context sensitive, in that a data gathered during a collection step can influence follow-on data collection requirements. When off-line, data collection is saved on the mobile device and asynchronously uploaded to the server database when a connection is subsequently established.

The server application will used by business managers and analysts. The server application provides a semantic database for the collected data and performs pipeline segment risk analysis. The risk analyses results is presented to the user through tabular displays and geographical visualizations.

The SmartForms Platform includes an environment for configuration of forms displayed on mobile devices, workflows and rules that used the data collection context to drive the sequence of forms or data displayed. Developers use the configuration environment to implement the use cases described in this document. Developers are familiar with proper use of tools in the configuration environment and create services that can be deployed in the SmartForms server, connect to necessary data sources and targets, and display the SmartForms in multiple field devices.

User Personas

Three user personas are identified for the SmartForms application; the field technician, the configuration developer, and the manager/business analyst.

a. Field Technician

The field technician is responsible for the collection of field data for a variety of field activities. Each field activity has a unique set of information that needs to be collected and stored into a central database.

The field technician needs to confirm that they have gathered all of the appropriate field information for a field activity, which may vary depending on the intermediate data gathering. For the POC application, it is assumed that the field technician is:

- Already trained and qualified for data collection for the defined field activities.

- Already familiar with the basics of using an Android smartphone app.

b. Manager/Business Analyst

The manager/business analyst is responsible for determining and prioritizing field activities and pipeline network maintenance based upon historical and current field data. They are also responsible for developing and maintaining the SmartForm work flows used by the field technicians for collecting field activity data.

The manager/business analyst wants to be sure that all required data for a field activity is gathered and that the new information gets uploading into the central database for analysis. Further, if a field technician gathers new information that warrants extra data collection, the manager/business analyst would like to interactively guide the field technician while they are still in the field.

The manager/business analyst want to be able to focus their maintenance and improvement capital on the pipeline network segment that are at the highest risk of failure. They will use the SmartForms application to perform analysis queries of the central database to produce ranked risk-based assessments of pipeline segments within pipeline network. For the POC application, it is assumed that the manager/business analyst is:

- Knowledgeable about how to interpret field data to determine risk of failure
- Knowledgeable about the process(es) that field technicians should use to collect data
- Minimal experience with creating programmatic process workflows

c. Configuration Developer

The configuration developer creates and updates the data collection workflows used by the field data collection mobile app. The developer uses business modelling tools to define and test data collection processes. For the POC application, it is assumed that the workflow developer is:

- Knowledgeable about the process(es) that field technicians should use to collect data
- Has experience with creating programmatic process workflows
- Knowledgeable about the relations and dependencies that influence risk of failure to a pipeline network
- Understands how to build and interpret Bayesian network models and apply them within the workflow

Specifications

Functionality

High-level requirements that are supported by the SmartForms applications include the following.

- Runs on an Android smart (mobile) device.
- Supports disconnected data collection in cases of no internet access.

- Supports syncing of disconnected data once internet access is restored.
- Supports flexible data entry so that any non-standard data can be captured.
- Support the auto-generation of a point location, this should represent the GNSS point where the user created the record. The point feature should be inserted into ESRI Geodatabase.
- Supports taking a geotagged photo that is retained and related to the record of interest.
- Supports User Authentication including "organization" or "domain"
- Support Deployment Configurations
- Manage Data Connection
- Requirement details to be elicited by the vendor as part of the project's execution.

User Interfaces

The field user interface does the following:

- Support basic Android style and functionality conventions.
- Support different screen sizes, different screen densities
- Implement adaptive UI flows
- Follow Android icon guidelines
- Use proper margins and padding
- Handle device orientation changes
- Use large, obvious tap targets (buttons, list items)

The application monitoring user interface has the following:

- Display:
    - The number of active sessions (devices that have connected to the server and have not disconnected)
    - The number of mobile devices currently connected to the central server
    - The number of mobile devices that connected to the server but are not currently connected (have become decoupled but have an active session)
- Display for each connected device
    - A visualization showing the state of the application on the mobile device. For example, if the application defines 5 screens/forms, which are displayed in some order (dependent on the application logic), then a visual that shows the form that was last displayed to the user on the mobile device.
    - Provide a historical record of device connectivity

The business manager/analyst user interface does the following:

- web-based
- Present a gains chart for the pipeline network
- Present a color-coded geospatial view of the network segment risks

The configuration developer user interface has the following features:

- web-based
- Provide graphical tools for viewing and modifying the field data collection forms and workflows
- Push updated and approved forms and workflows to field mobile devices.

Performance

Application simultaneously supports, without significant response time impact:

- 25 mobile device users
- 1 administrative (configuration) user
- 2 management/business analyst users

Application responds to mobile device user requests within:

- 5 seconds for requests requiring on-line connectivity
- 3 seconds for off-line mode requests

Application availability is:

- During normal business hours as specified for the use cases
- Field device geographic coverage is to limited to area(s) specified by detailed design use cases

The deployed POC application uses local mobile devices loosely-coupled to a central, cloud-based server. The POC demonstrates mechanics that permit the mobile device application to operate properly when the server connection is unavailable, and for synchronization to occur when connectivity to the server is restored.

The mobile application must run within the available RAM and flash memory on the mobile device. The mobile application must function in both on-line (connected to a remote server) and off-line (disconnected from remote server) modes. When off-line, the application provides local (on-device) functions as defined in the use cases. Whenever an off-line device reestablishes a connection to the remove server, local data is automatically synchronized with the remote.

The application must accommodate as much as possible network latencies that limit responsiveness of distant servers providing information to the mobile application.

Operations & Maintenance

Users must download and regularly update the mobile application. Data stored locally on the mobile device must be backed up by the user using 3rd party backup software that is supported by the SmartForms application.

SmartForms application sessions that remain open on the server for more than 1 hour after connectivity is lost is automatically closed and users must re-start the application to begin a new session. Sessions that become disconnected from the server (decoupled) for any period of time less than 1 hour must be synchronized with the server upon reconnection. No provision for re-initializing cancelled sessions is supported.

Systems Integration

The SmartForms application requires two systems interfaces; a web services interface to support communication between the SmartForms remote, mobile device application and Microsoft Internet Information Services (IIS) web server, and an Azure Service Bus messaging interface between all database servers, and Microsoft Azure ecosystem services/applications, including GIS.

All communication between the remote SmartForms application on Android and the SmartForms server application must use web services protocols. Details on IIS Web Development can be found at https://msdn.microsoft.com/en-us/library/ms525945(v=vs.90).aspx

All communications between ESRI servers, Azure ecosystems services/applications, database applications and the SmartForms server application is via JSON-LD messages across an Azure Service Bus.

Details on use of the Azure Service Bus can be found at https://azure.microsoft.com/en-us/documentation/services/service-bus/

JSON-LD information can be found at http://json-ld.org/

Security

The SmartForms applications must support communication protocols and provide suitable monitoring functions for protecting against malicious or accidental access, modification, disclosure, destruction, or misuse. Security functionality includes:

- Session password-based sign-on & authentication with configurable authorizations/permissions based on user role
- Single factor authentication (SFA) is used. Future versions are expected to require improvements in authentication. Therefore, the SFA implementation must not limit conversion to improved authentication functions in the future.
- Use of https for web services calls
- Support for existing mobile device security systems and softwares
- Server-side components uses suitable firewall software together with activity logging that enables forensic analysis of security issues

Data Management

The SmartForms application supports persistent storage of:

- Use case configurations
  - Activities and rules
  - Use case templates
- GIS data
- Collected field data

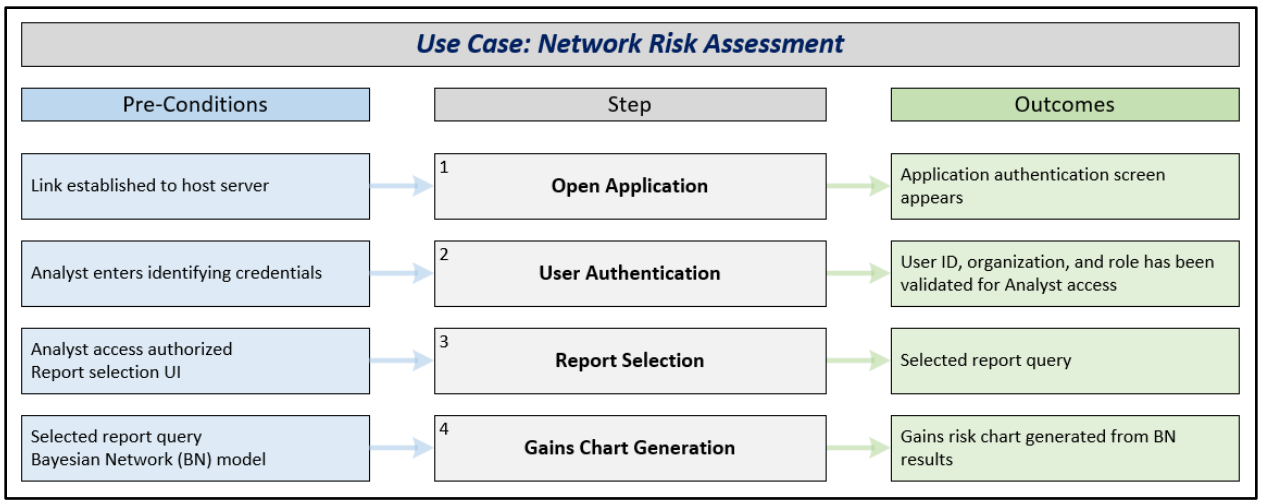
SmartForms data management functionality includes, as required for the use cases defined in this document:

- predefined OWL 2.0 ontologies for classes, attributes (data properties) and relations (object properties)
- temporal (time stamp) management at the required level of time granularity for
  - data collection
  - data history
- database assertion and access rates is lower than network and application latencies such that database activity does limit application performance

Use Cases

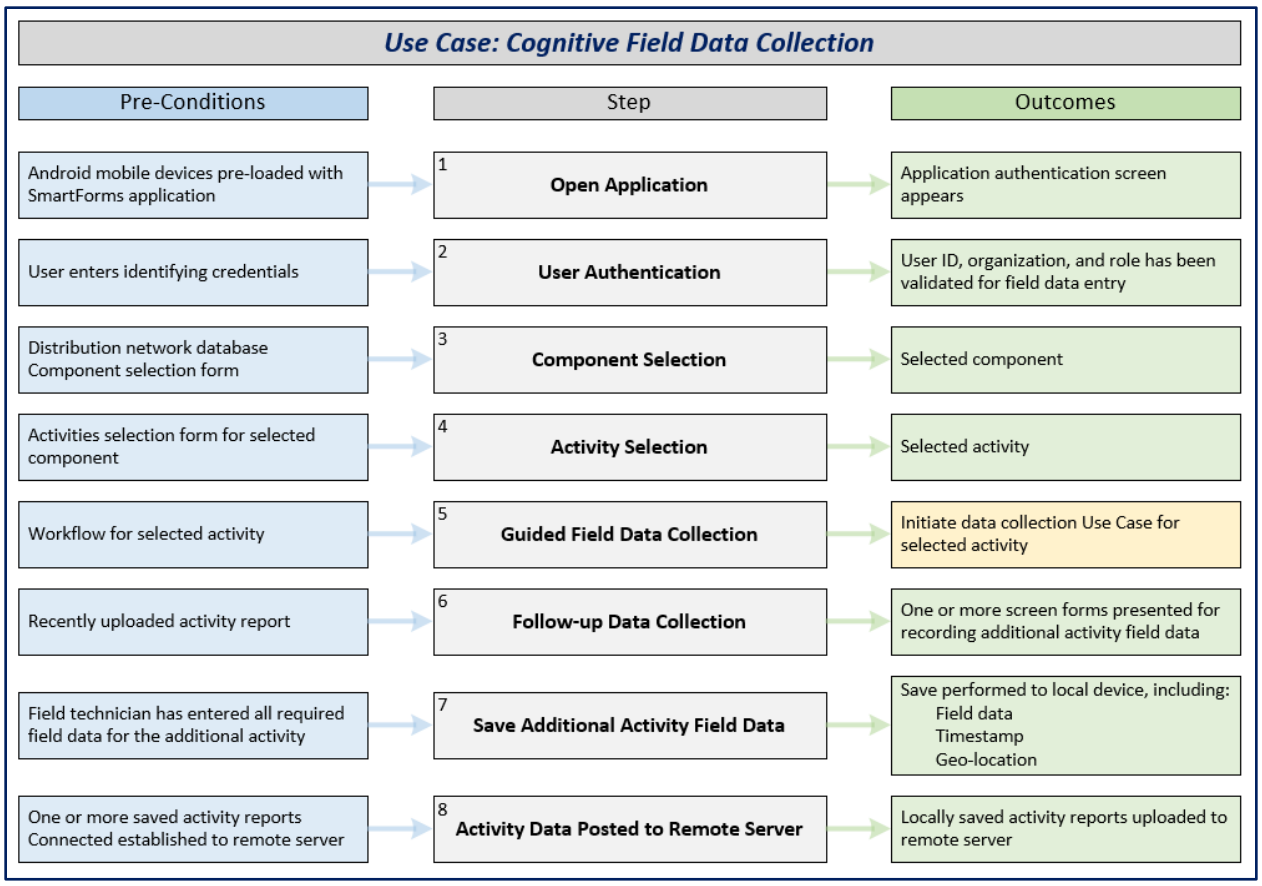
Network Risk Assessment

Goal	Perform Bayesian risk assessment for a pipeline network using current knowledge base
Pre-Conditions	Link to SmartForms host server User authentication with analyst access
Successful End	Risk report and gains risk chart generated for selected network
Abnormal End	Unable to generate report due to programmatic exception or missing information
Agent(s)	Business Manager/Analyst SmartForms host server
Trigger	Ad hoc request by a Business Manager/Analyst



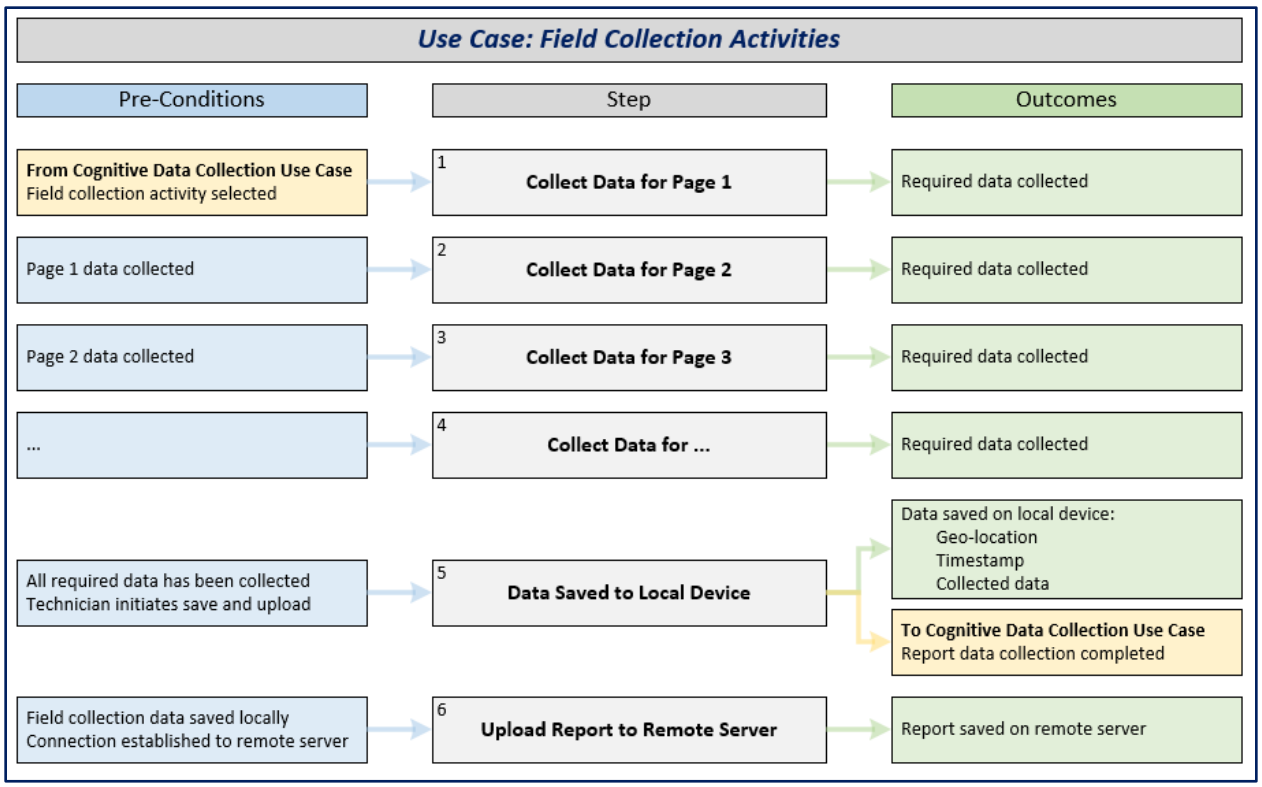
Cognitive Field Data Collection

Goal	Provided guided pipeline network field data collection via an Android mobile device
Pre-Conditions	Android mobile device with SmartForms application loaded Qualified Field Technician with user authentication for technician access
Successful End	Completed activity report uploaded to host server
Abnormal End	Activity report information not received by host server Incorrect and/or incomplete information collected and saved for selected activity
Agent(s)	Qualified Field Technician SmartForms mobile application SmartForms host server
Trigger	Ad hoc initiation by Qualified Field Technician



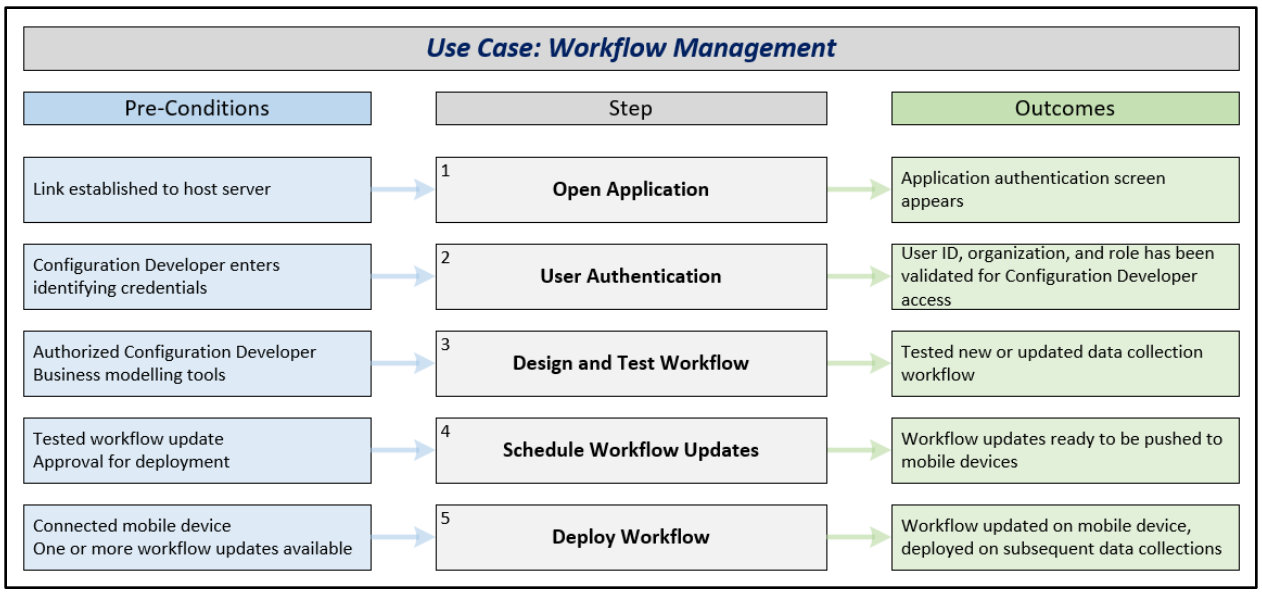
Field Collection Activities Use Case

Goal	Collect specific field data for a field activity, including: • Manual entries • Scanned barcodes • Photographs
Pre-Conditions	Qualified Field Technician has selected a field activity in the SmartForms mobile application
Successful End	Completed collection of field data
Abnormal End	Application fails to present appropriate forms Application collects incorrect data for activity Application fails to record governance data
Agent(s)	Qualified Field Technician SmartForms mobile application SmartForms host server
Trigger	Selection of a field activity within the SmartForm mobile application



Workflow Management

Goal	Create or update field data collection process workflow in the mobile app
Pre-Conditions	Link to SmartForms host server User authentication with Configuration Developer access
Successful End	Workflow created or updated and available to SmartForm mobile app as an update
Abnormal End	Unable to create or update workflow successfully Unable to download the latest workflow into mobile app
Agent(s)	Configuration Developer SmartForms host server
Trigger	Ad-hoc by a Configuration Developer



Design

This section describes the design of a proof-of-concept (POC) SmartForms application, sufficient to enable developers to plan and begin work on creating the POC application. The design comprises two sections that follow this introduction;

1. the Configuration Management section, describing the design of the environment that is used to manage the lifecycle of the SmartForms specified in the 7 required and, by extension other, use cases, and
2. the Situational Reasoning section, describing how intelligence needed to drive SmartForms displays in mobile devices is implemented;
    a. interface with, baseline and store SmartForm component, activity and geospatial data in a knowledgebase,
    b. manage multiple models used in reasoning, and
    c. make data available to a user interface capability for visualization.

This document is updated as the project proceeds to include the latest understanding of the SmartForms application design. The version record contains the document version number, authors, publish date and notes for each version.

Configuration Management

This section describes the design of a Configuration Manager that enables definition of;

1. Forms or screens, implemented as HTML5 web pages, that contain user interface fields;
2. Workflows comprising a sequence of tasks and decision points that direct the display of Forms based on data entered into the user interface fields.

The capabilities of the Configuration Manager implement a generic ability to define a finite state machine[10] that determines the states and sequence in which a specific series of Forms are displayed during a field data collection session **Figure 7-25**. In creating Forms, Developers effectively specify States, Data Items, Methods and Workflows that implement the Form state transitions in the 7 required project use cases, or new use cases defined in the future.

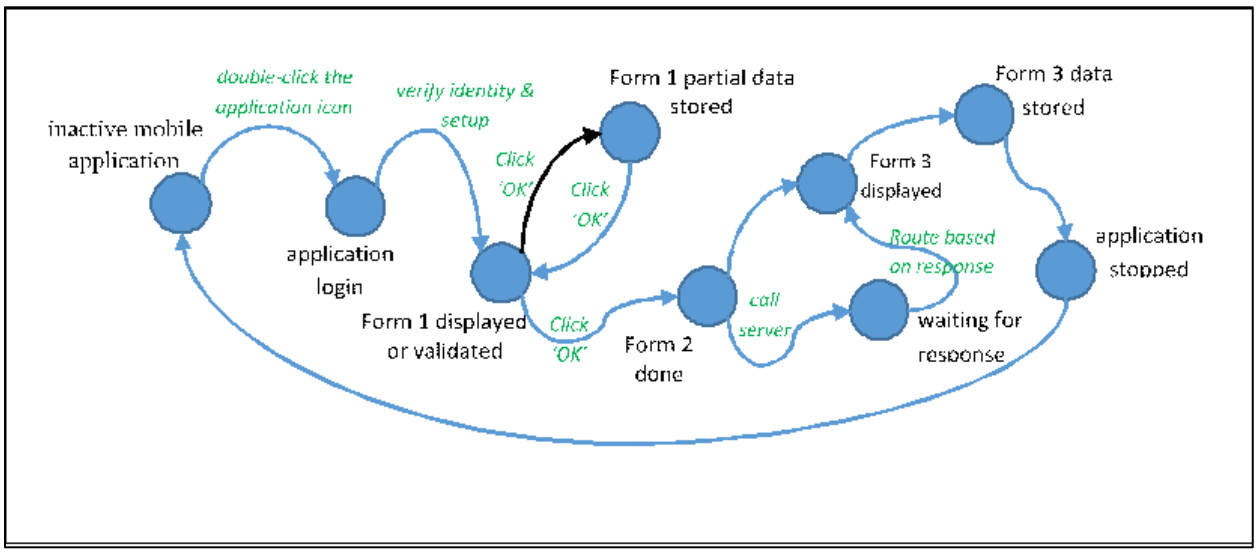


Figure 7-25. Example sequence of states in a SmartForms application, effectively configured in the Configuration Manager

Situational Reasoning

This section describes design of the SmartForm manager-based services for;

1. Representing situations that describe relevant states or conditions that may exist in data collected by the SmartForms mobile application used by Field Technicians;
2. Representing Bayesian network structure and data and relations to situational representation in order to allow for Bayesian network node values to inform situational determination;
3. Representing situation action networks that define the way in which goals are achieved by SmartForm manager agents.

SmartForm manager provides situation centric cognitive- and knowledge-based ontological services involving a variety of numeric and logical reasoning techniques. These capabilities can be packaged in a variety of ways ranging from small components deployed within Azure functions up to large systems distributed over an elastic pool of virtual machines.

---

10 The states and their transitions can be represented in a state chart diagram.

Composability is a strong theme throughout the architecture, allowing packaging decisions to be made based on what scalability requirements are required for different parts of the solution.

In this application, SmartForm manager is primarily a consumer of data generated by the SmartForms workflows. The data is important to support the continuous improvement of models from which risk-based decisions are made. Activity logging supports the provenance of SmartForm manager data, so with this in mind the primary objectives for SmartForm manager are as follows:

1) To be the system of record for provenance data, including activity logging records.
2) To manage an interface with an ESRI system so that geographical visualization of data is possible. This data is a mixture of slowly changing master data (configuration of the gas network) and day-to-day information (including photographs) from the field data collection sessions.

To provide an interface with business intelligence software so that gains charts and other data visualizations can be produced.

Component Architecture

Figure 7-26 illustrates the principal components and the relationships between them. The opportunities to use communication techniques described above are shown:

- Enterprise Service Bus messaging is used to convey information from the K2 Workflow environment to SmartForm manager.
- SmartForm manager injects data into the ESRI ArcGIS system using a RESTful API
- ArcGIS can also request extra web content from SmartForm manager as needed.

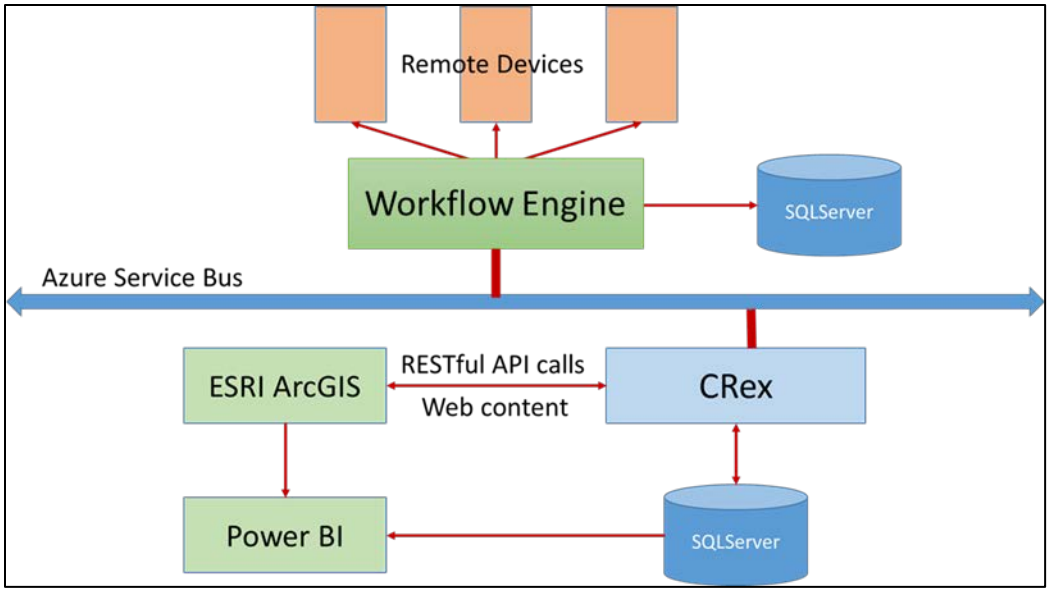


Figure 7-26. Example sequence of states in a SmartForms application, effectively configured in the Configuration Manager.

Model Management Techniques

Anything that has information associated with it can be represented as an Entity in SmartForm manager. Elements of models can be entities, as can real-world objects, events, people, other agents and the activities they perform. All entities have unique identifiers. Simple entities can be composed into more complex entities, so a model composed of model elements can be an entity, as can departments composed of people, roles and responsibilities, and plans and workflows can be entities composed of activities. A situation is an entity composed of objects, people and agents plus a narrative of events and activities that involve them as time passes.

If the information about an entity changes over time, SmartForm manager tracks those changes by maintaining a version history of each entity. Versions of entities are identified by a combination of the entity identifier and the identifier of the activity that created the new version. Some entities are expected to remain the same across time; the main examples of these are entities that represent events and entities that represent messages exchanged between parts of the system.

Sometimes, SmartForm manager needs to reason about different ways in which the world might change. To keep track of different possible futures, the version history can have branches. Each branch is also given an identifier.

Activity Logging and Provenance

To keep an accurate record of activities within the system, an activity model is required. In addition to activities, the primary entities in this model are agents (software agents, organizations, and people) and the data entities that are produced and consumed by the

activities. This combination of entity types not only serves as an activity log, but also provides rigorous provenance for data entities produced by the system.
Auditability of activity logs is achieved by maintaining two kinds of model in SmartForm manager.

1) A model derived from designed workflows that captures how a data collection should proceed.
2) A model of what actually happened during a field data collection session.

It is then possible to assist the auditing process by identifying significant differences between the desired and actual data collection.
Each of these models is represented by a situation. The setting of the situation's narrative includes a gas network segment, a field technician along with a location and a time interval over which data collection occurs. In the desired model, these are represented by place-holders whereas in the model of actual data collection these are replaced by references to real pipes, people etc. This allows the same model elements to be used for both models.

Data Management

Data managed by the SmartForms application includes;

- Workflow configurations & templates
- Form configurations & templates
- Pipe networks data including;
    - ID = Unique id for each AldylA distribution mains
    - End1_Longitude, End1_Latitude = Longitude and latitude of one end of the mains
    - End2_Longitude, End2_Latitude = Longitude and latitude of other end of the mains
    - Area_Code = 1 to 5
    - Year_Installed = Year In which the mains was installed (1969 - 1980)
    - Years_In_Service = Current year - Year_Installed
    - Probability_of_Failure_5yrs = Probability of failure in 5 years
    - Probability_of_Failure_10yrs = Probability of failure in 10 years
- Collected field data
- For each of the above, a timestamp and validity interval

Predefined OWL 2.0 ontologies are provided for classes of objects, their attributes (data properties) and relations (object properties), and elements required for situational and Bayesian network concepts.
The storage technology is Microsoft Azure SQL Server.

Aldyl-A Pipeline Network Model

Information about the gas network is represented as a composable, layered model. The most basic layer contains the geographic representation of each network segment. Other layers add more data so that views combining several kinds of data can be composed easily **Figure 7-27**. The unit of composability is the amount of knowledge for one network segment for one layer, and these is represented as versioned entities in SmartForm manager.

The set of layers required for this model are as follows:

1) Geographic layer. Provides the geographical point data needed to locate the network segment. This layer also contains details of features near the pipeline segment that combine with risk-of-failure to generate a risk metric.
2) Physical composition layer. Provides data on the dimensions and material of the pipe used in the network segment, along with estimates of replacement cost for the segment.
3) Physical provenance layer. Provides details of the manufacturer, data of manufacture and date of deployment of the pipe. It also provides a maintenance history of the pipe since deployment that includes maintenance techniques used.
4) Risk analysis layer. Contains the results of Bayesian analysis of risk using the Aldyl-A model.
5) Custom layers. As needed to support the kinds of information made available from field collection workflows.

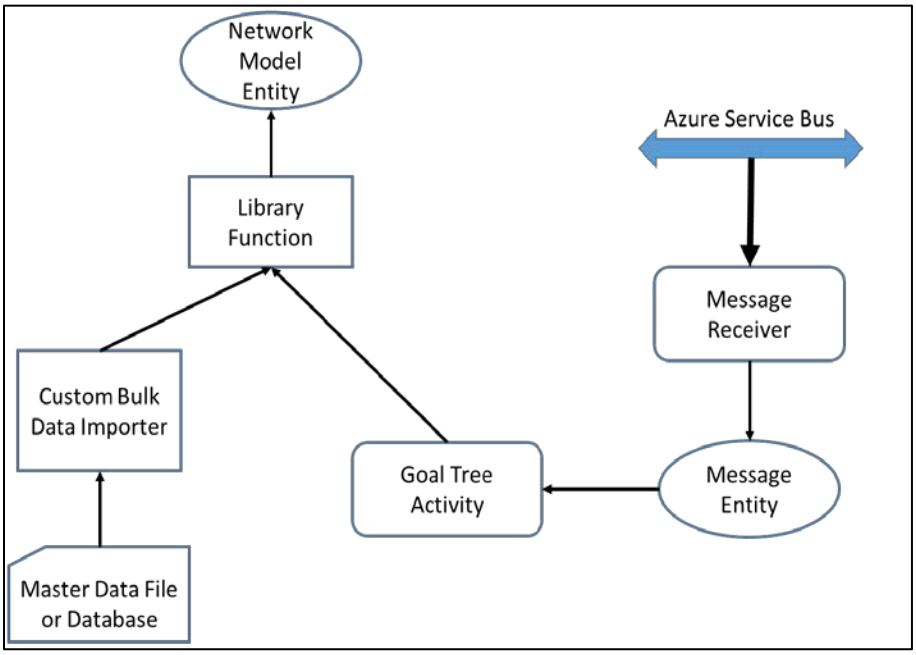


Figure 7-27. Layer population techniques

The two main pathways for data to arrive in the network model are the bulk import of master data from external resources (data files, databases etc.) and the translation of messages that arrive in SmartForm manager. Noting that revisions to master data records may result from messages delivered to SmartForm manager, it makes sense that there is a library of functions for creating master data elements that are used by both techniques as shown below. The primary difference between the information that is stored in each case is in the provenance data that is created.

User Interface and Integration with ESRI/ArcGIS

SmartForm manager data can be made available in ArcGIS through the RESTful API for the product. The required steps are as follows:

1) Data is made available in records. A schema for the records must be declared. Once this has been done, the schema is used for all records of that type.
2) A target for a new data record is created or chosen in the ArcGIS data. A target might be a geographical point created in ArcGIS to represent the location of a field engineer during data collection, a geographical point created to represent the location a photograph was taken, or an extended geometric feature such as the representation of the pipe segment.
3) A data record is created and populated using data values extracted from one (or more) levels in the network model, typically for a single network segment.
4) The data record is associated with the selected target in ArcGIS.

New records created in this way become viewable through the ArcGIS user interfaces and the data is available for queries made against ArcGIS data stores.

Not all data that is accessible through the ArcGIS user interfaces must be copied into the ArcGIS database. Data can also be made available by storing a URL in an ArcGIS record. This is the method used to associate photographs with geographical points since the images are stored in Azure Blob Storage and the URL supplied to SmartForm manager. URLs can also be used to address knowledge in SmartForm manager, reducing the amount of data that must be replicated between the systems.

The primary means for a business analyst to interact with the system is through Microsoft Power BI.

Geospatial View

This solution takes advantage of standard components that already exist for projecting ArcGIS data into Power BI.

Gains Chart

A gains chart component looking like the following has been proposed **Figure 7-28**

Probability of Failure within 5 years	Number of Segments	Percent of Segments	Cumulative %	Average Years in Service	Linear Feet	Replacement Cost	Cumulative Replacement Cost
90%+	183	2.38%	2.38%	51.8	12,426	$2,795,756	$2,795,756
80-90%	975	12.66%	15.03%	44.1	82,997	$18,674,363	$21,470,118
70-80%	985	12.79%	27.82%	44.0	81,316	$18,296,065	$39,766,183
60-70%	993	12.89%	40.71%	44.1	88,431	$19,896,967	$59,663,150
50-60%	398	5.17%	45.88%	40.4	31,286	$7,039,312	$66,702,463
40-50%	218	2.83%	48.71%	44.1	16,707	$3,759,120	$70,461,583
30-40%	64	0.83%	49.54%	43.9	4,141	$931,670	$71,393,253
20-30%	1,162	15.09%	64.62%	45.3	121,652	$27,371,634	$98,764,887
10-20%	1,494	19.40%	84.02%	37.0	131,774	$29,649,092	$128,413,979
<10%	1,262	16.38%	100.40%	40.7	123,825	$27,860,680	$156,274,660

Figure 7-28. Example of a Gains Chart

The aggregate data in this table can be computed by Power BI based on a view of data extracted from various layers of the network for each segment of the network **Table 7-1**

Table 7-1. Data for creation of gains charts

Field	Source
Probability of Failure	Produced during Bayesian analysis with the Aldyl-A model; stored in the risk analysis layer of the network model.
Years in Service	Stored in the physical provenance layer of the network model.
Linear Feet	Stored in the geographic layer of the network model
Replacement Cost	Stored in the physical composition layer of the network model.

SmartForm Ontologies

Foundation - The foundation ontology is a SmartCloud ontology® defining the modeling elements for composing settings and narratives using Situation Theory.
Provenance - This is a small SmartCloud ontology® that places the design patterns from PROV-O (see below) within the context of situations as represented using the foundation ontology.
Standard and 3rd Party Ontologies
Units of Measure - The OM-1 ontology focuses on units of measure, their dimensionality and the quantities and measurements that use them. (see http://www.wurvoc.org/vocabularies/om-1.6/)
Organization - The Organization Ontology is a W3C Recommendation since 16th January 2014. It is centered on terminology that allows the modeling of organizational structures with important connections to the Roles and Agents required by the PROV-O ontology. (see http://www.w3.org/TR/vocab-org/)

People - The Friend-of-a-Friend (foaf) ontology provides terminology about people, connections between people, and knowledge related to people. (see http://xmlns.com/foaf/spec/)
Media - The Dublin Core ontology is a small, frequently used ontology for describing media items in the semantic web. (see http://wiki.dublincore.org/index.php/User_Guide)
Provenance - The PROV-O ontology is a W3C recommended ontology for representing the use and production of data entities during activities performed by agents. It is widely used within the semantic web community and represents state-of-the-art for activity logging and provenance. (see http://www.w3.org/TR/prov-o/)

Keyhole Inspection Form, First Response Form, Audit Report Form

The Keyhole Inspection or First Response form improves data collection process by guiding operators through a series of relevant questions during inspection of the pipeline segment being inspected. The questionnaire is intelligently generated through decision-driven workflows where the data entered by operators determines the next question for data entry. The SmartForm runs on an android mobile device and supports flexible data entry so that any non-standard data may be captured. The geolocation of the inspection is automatically captured along with other details. **Table 7-2** lists the details of Keyhole Inspection and First Response forms.

Table 7-2. Keyhole Inspection Form and First Response Form

Smart Form type	Keyhole Inspection Form or First Response Form
Goal in Context	Create, view and edit a record for keyhole data gathering
Scope & Level	Scope = Intelligence Field Data Collection. Level = User Goal
Preconditions	The user should be familiar with the procedures and recommended practices to collect keyhole or first response data.
Success End Situation	The user successfully enters the keyhole or first response information (the mobile application works even when the internet connectivity is lost). The mobile application posts the data collected to the remote server (if there was no internet during data collection process, the data is posted to the remote server once the application is online).
Failed End Situation	Inability to operate in offline mode. Inability to save and post the data collected to remote server.
Primary Agent Secondary Agent	Field Operator Qualified Field Personnel
Trigger	A Field Operator (or any qualified field personnel) is out in the field to gather pipeline data.

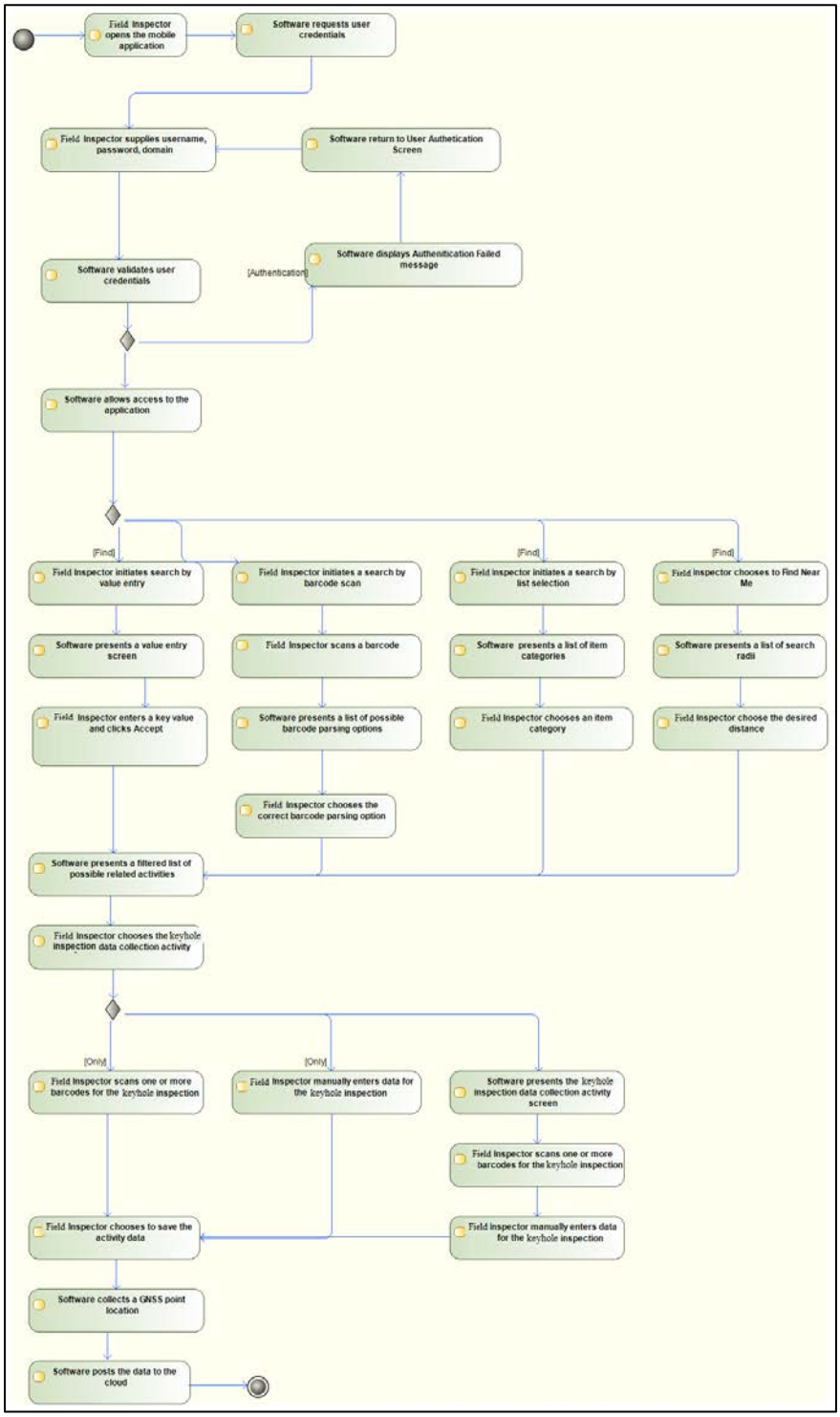
Smart Form type	Keyhole Inspection Form or First Response Form	
DESCRIPTION	Step	Actions Leading to the Successful Goal
	1.	Field Operator opens the mobile application
	2.	Field Operator enters user credentials in the login screen
	3.	Field Operator initiates search of the gas components (e.g., pipeline, valves, etc) by: - Value entry - Barcode scan - List selection - Find near me (based on location)
	4.	The mobile application presents a filtered list of possible related activities.
	5.	Field Operator chooses the keyhole data gathering or first response activity.
	6.	Field Operator navigates through a sequence of data entry screens to collect all relevant information. Each step/screen has one or more of the following: - Scans one or more barcodes - Manually enters data - Takes photos
	7.	Field Operator saves the activity data. The mobile application adds time and geo-location point information to the activity data.
	8.	The mobile application posts the data to remote data server.

RELATED INFORMATION	Keyhole Data Gathering; First Response
Priority:	1
Performance	Normal mobile application responsiveness is sufficient.
Communication channels between agents	Keyhole data/First response records on Android mobile phone.
Superordinates	Create, View, Edit records

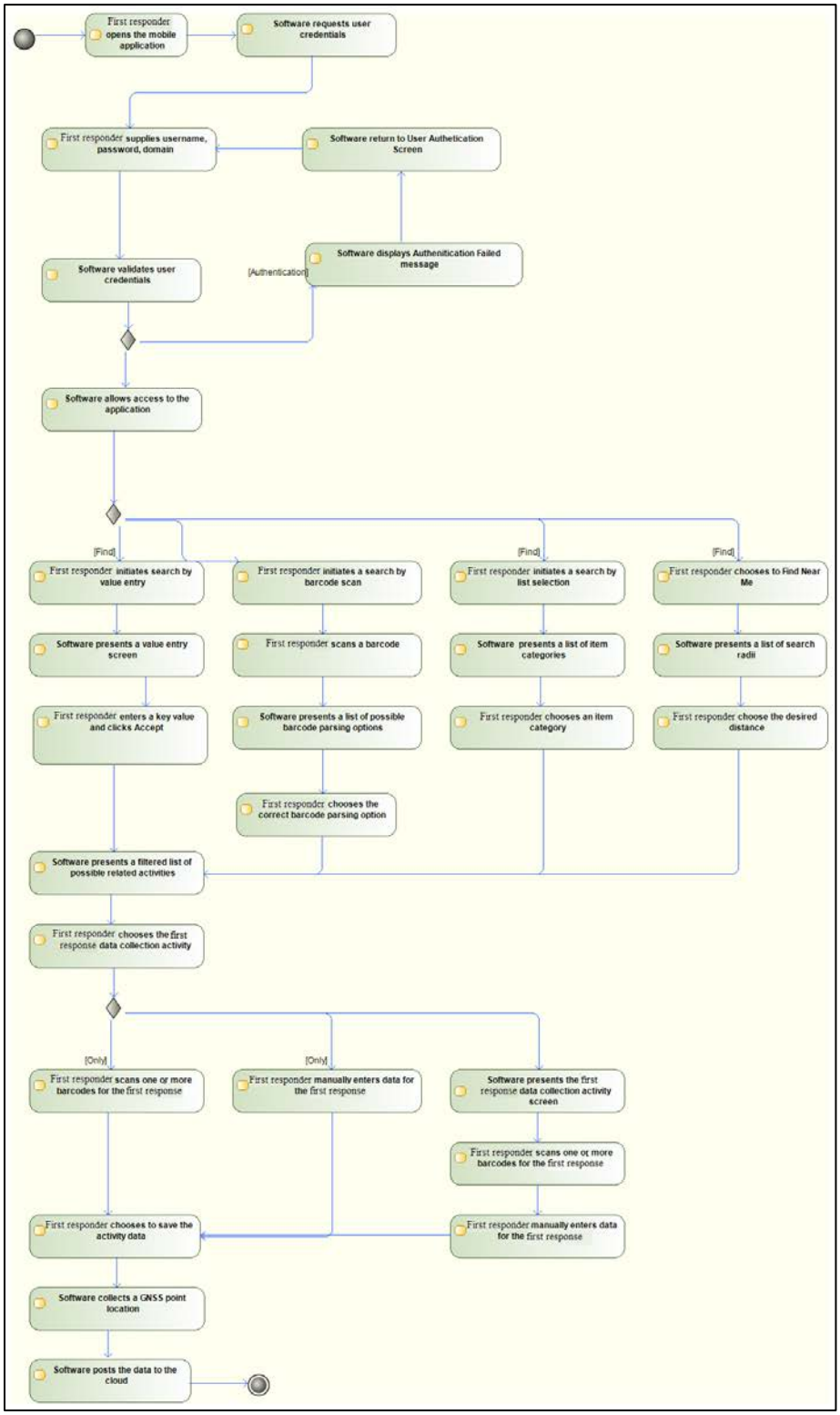
Gas operators are required to submit annual audit reports and incident reports to PHMSA. It takes quite an effort to collect all required information in one place to fill up the forms.
One of the goals of SmartForm is to replace the current manual report filing process and to automate the audit report generation by pulling together information from databases distributed across different business units. The audit report form helps to complete PHMSA F 7100.1-1 annual report form and the process audit prescribed in ASME B31.8S standard.

Activity Diagrams

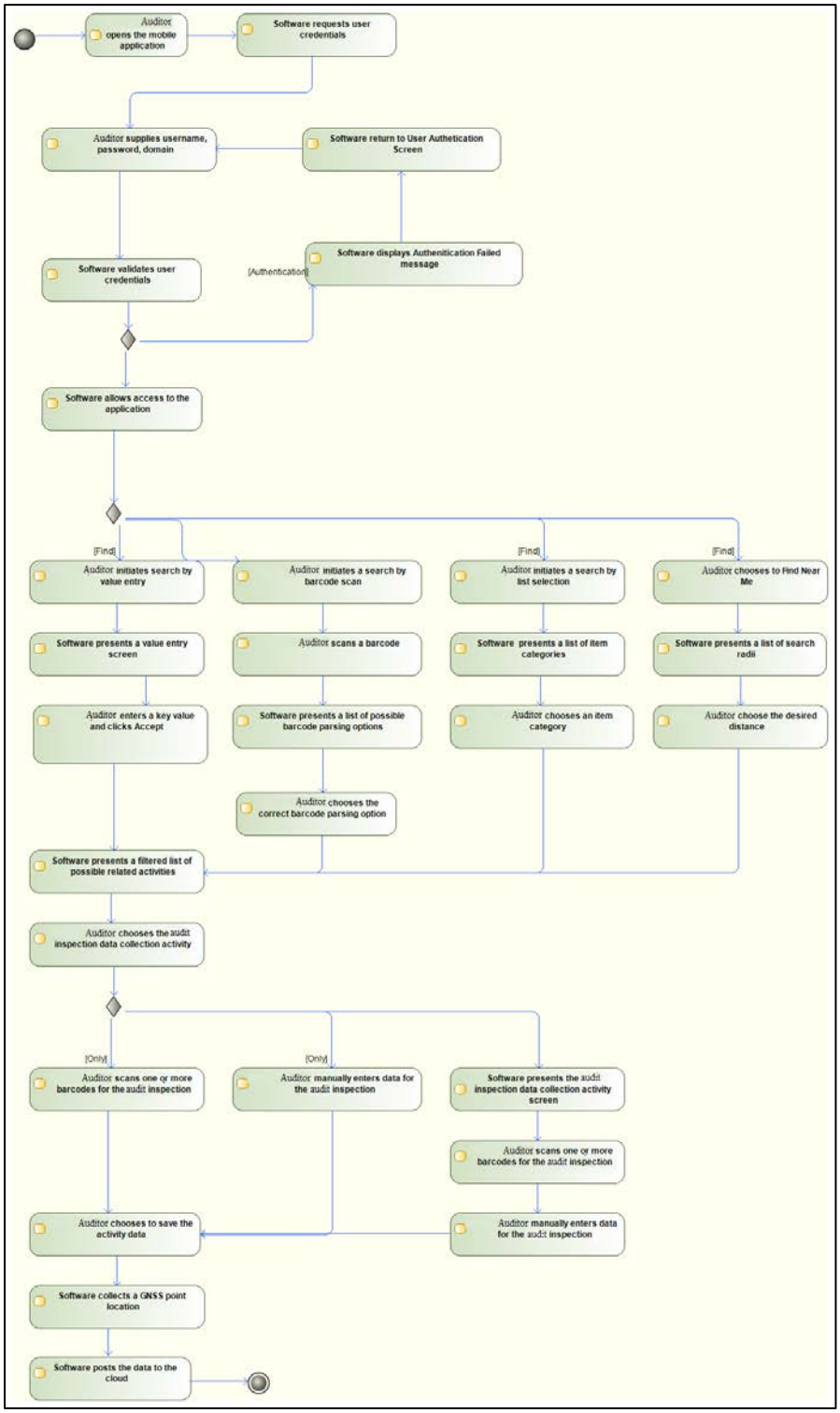
Activity Diagram – Create a tracking record for keyhole data gathering inspections



Activity Diagram – Create a tracking record for first response data gathering



Activity Diagram – Create a tracking record for audit event data gathering



Appendix A- Use Case: EDSS with SmartForms Demo

Pre-Conditions	Step	Outcomes
Laptop with server application started Server application loaded with demo network Mobile device with loaded app and connection to laptop server Demo BA logon credentials	1 **Business Analyst Logon**	User logged on and role authenticated ?Option menu displayed?
BN trained for demo pipeline network in developer UI	2 **Present Bayesian Network (BN)**	Viewer understands how BN is used to determine pipeline segment risk
User logged on with BA authentication Network Risk Analysis requested	3 **Display Network Risk Analysis Table**	Prioritized risk analysis table (gains chart)
Geographical Network Risk display requested	4 **Display ARCGIS Risk Visualization**	ARCGIS network visualization of pipeline segment risk generated
ARCGIS risk display	5 **Zoom to Demo Pipe Segments**	Pipe segments "scheduled" for field reports displayed
BPMN diagram for activity report that requires risk-based branching of the field data collection workflow	6 **Present Risk-Based Collection Design**	Viewer understands how risk-based data collection is to be implemented
Implemented activity report workflow that uses risk-based field data collection workflow branching	7 **Present Risk-Based Collection Workflow**	Viewer understands how risk-based data collection has been implemented on mobile device
Mobile device with loaded app and connection to laptop server	8 **Field Tech Logon**	User logged on and role authenticated Component query menu displayed
User logged on with Field Tech authentication	9 **Component Selection by List**	Demo component selected Report activity menu displayed
Demo component selected	10 **Activity Selection**	Activity report selected
Demo component selected Activity report selected	11 **High Risk Component Assessment**	Risk information for workflow guidance downloaded to mobile app Form 1 of report workflow displayed
Form 1 of selected activity displayed	12 **SmartForm-Led Data Collection**	TO BE EXPANDED All data collected for selected report
Completed Activity Report	13 **Data Uploaded to Host Server**	Report received and store by host server application
User logged on with Field Tech authentication	14 **Component Selection by List**	Demo component selected Report activity menu displayed
Demo component selected	15 **Activity Selection**	Activity report selected
Demo component selected Activity report selected	16 **Low Risk Component Assessment**	Risk information for workflow guidance downloaded to mobile app Form 1 of report workflow displayed
Form 1 of selected activity displayed	17 **SmartForm-Led Data Collection**	TO BE EXPANDED All data collected for selected report
Completed Activity Report	18 **Data Uploaded to Host Server**	Report received and store by host server application
New Activity Report(s) received by host server application	19 **Situational Evaluation**	New information is available for risk analysis queries Situational HCA level report issued
User logged on with BA authentication Network Risk Analysis requested New activity report information	20 **Display Network Risk Analysis Table**	Prioritized risk analysis table (gains chart)
Geographical Network Risk display requested	21 **Display ARCGIS Risk Visualization**	ARCGIS network visualization of pipeline segment risk generated
ARCGIS risk display	22 **Zoom to Demo Pipe Segments**	Pipe segment that received new field report data displayed Status updated based on latest data

Deployment and Testing

In this section, we describe our deployment and testing methodology, test data, test cases, and the various types of testing performed on Keyhole inspection form, first response form, audit form, and Bayesian network.

Table 7-3. Technical Specification of the Deployment Environment

	Details
Vendor	Microsoft Azure
Operating System	Microsoft Windows 10 Virtual Machine
Virtual Machine Hardware	2.4 GHz Intel Xeon® E5-2673 v3 (Haswell) processor
Processor	2 cores
RAM	14 GB
Disk Space	100 GB

Test Data Preparation

To test the system, a synthetic virtual gas pipeline network of Aldyl-A plastic pipes was prepared in GIS system. The pipe network consists of 7706 piping segments divided into five regions. Each segment was characterized with lab and synthetic data describing various pipe attributes such as pressure, material, installation date, repair date, associated leaks, etc. **Figure 7-29** shows the map of the test pipeline network.

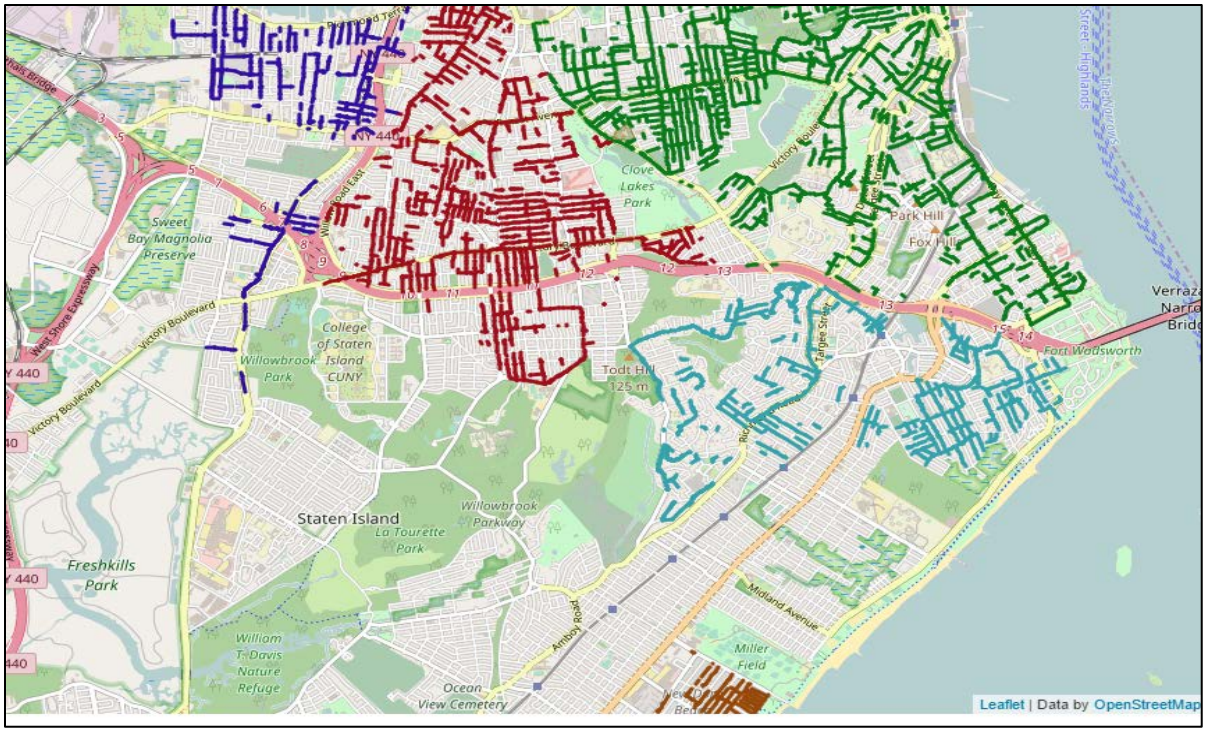


Figure 7-29. Test gas distribution network

SmartForms such as keyhole inspection and first response are connected to the test gas network and serve as a field data collection tool for individual pipe segment. The Bayesian network is also connected to the test gas distribution network and is used to calculate the mean life prediction and the risk of failure for every pipe segment. The testing of risk analysis can be conducted at segment level or region level.

Test Case Preparation

Test cases are essential tools for systematic planning and execution of testing scenarios. In this project, a number of test cases were prepared to list a set of conditions to determine whether the deployed probabilistic decision support system and SmartForms satisfies the requirements and functions correctly. The system was tested and refined until all test cases were passed. The specification format for test cases used in this project is listed in the table below.

Attributes	Definition
Test Case ID	Unique ID of the test case
Test Case Summary	Goal of the test case
Prerequisites	Any pre-condition that must be met before executing the test
Test Procedure	Detailed steps to execute the test
Test Data	Data to use for testing

Expected Result	Expected result of the test
Actual Result	Actual result from the test
Status	Pass/Fail to indicate the test result
Remarks	Additional comments about the test
Created By	Person who created the test case
Date of Creation	Date of test case creation
Executed By	Person who ran the test
Date of Execution	Date of test execution
Test Environment	Hardware and software environment in which the test was performed

Testing of Keyhole Inspection Form

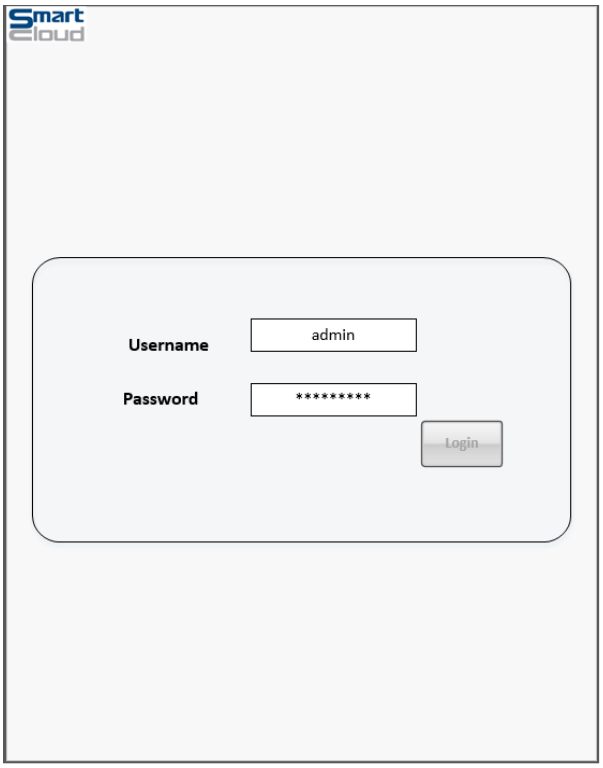
Functional Testing

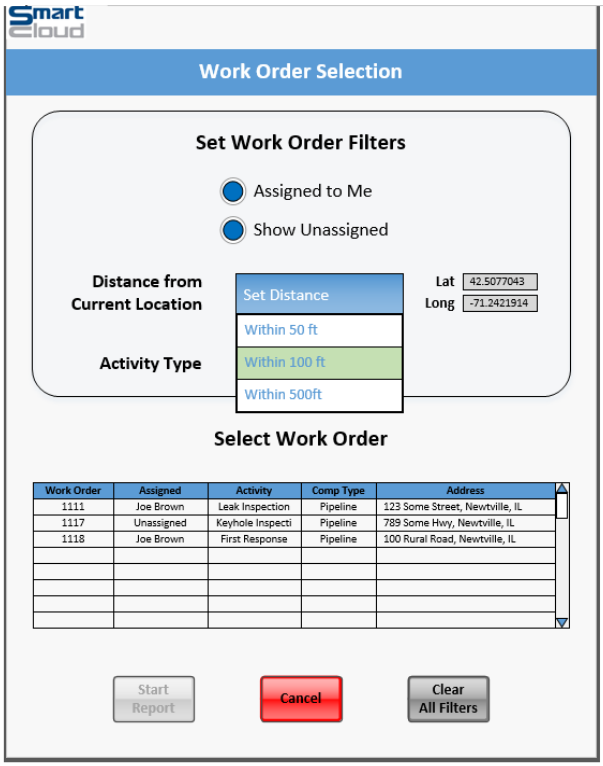
Each functionality of the Keyhole inspection form was tested to verify that it meets the requirement specification. An appropriate input is entered in the system and the output is verified by comparing the actual result with the expected result. The table below illustrates the steps taken for functional testing and lists the pass/fail grade for each testing step.

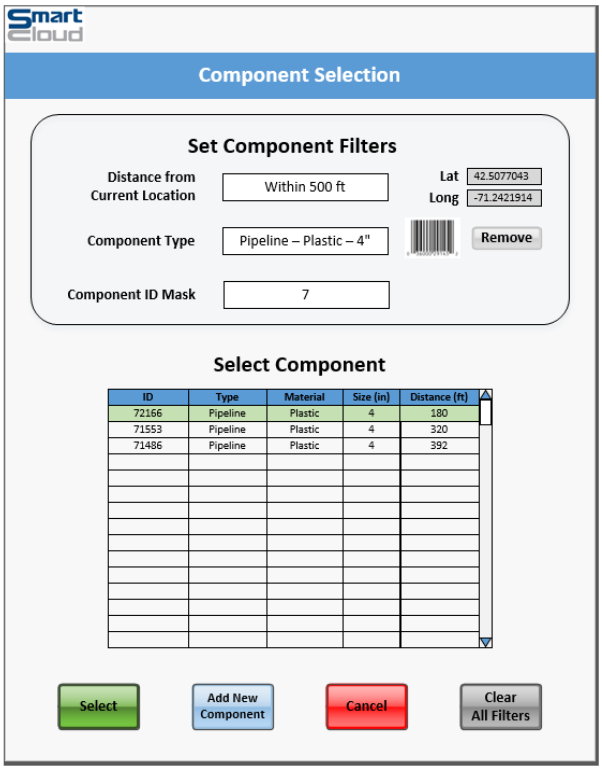
Step #	Step Description	Test Data	Expected Results	Test Pass/Fail
1	Enter login information	username: *admin* password: *Admin12248*	Show keyhole inspection form main page	Pass
2.a	Filter existing pipelines by barcode	ASTM F2897 barcode	List pipelines matching information in barcode	Pass
2.b	Filter existing pipelines by distance	100 feet	List pipelines within 100 feet of current location	Pass
2.c	Filter existing pipelines by type	Aldyl-A plastic pipe	List all Aldyl-A pipelines	Pass
3.a	Select existing pipeline from filtered list of pipelines	Pipeline with ID 72166	Show next screen to view historical data of the pipeline selected	Pass

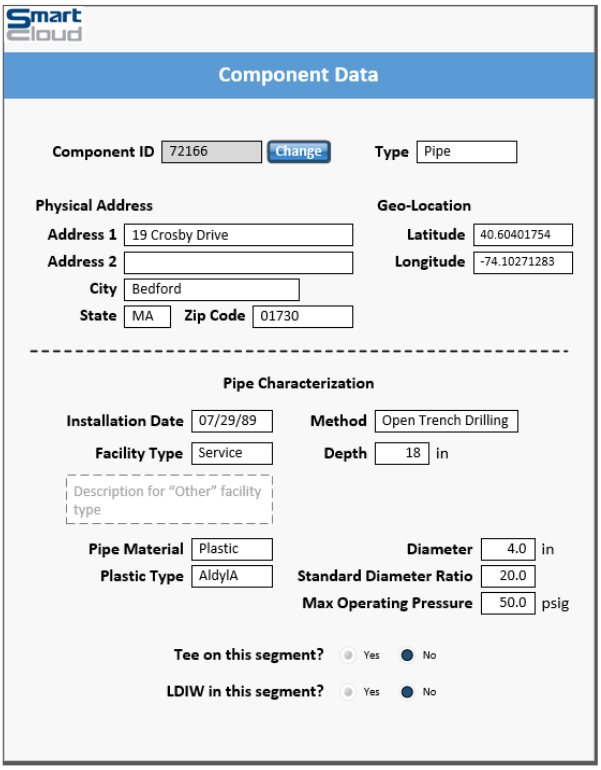
3.b	Add new pipeline information	Pipeline with ID 77667	Save new pipeline information and show screen to enter data	Pass
4	View historical pipeline data		Show past data and allow updates	Pass
5	Enter new keyhole inspection data		Collects all keyhole inspection data. As a next step, enable Save button	Pass
6	Save the entered keyhole inspection data	Press Save button	Save the data in cloud database	Pass

Sample screen shots of the functional testing are shown in figures below.


Smart Cloud
Username
admin
Password
****
Login


Smart Cloud
Work Order Selection
Set Work Order Filters
Assigned to Me
Show Unassigned
Distance from Current Location
Set Distance
Within 50 ft
Within 100 ft
Within 500ft
Activity Type
Lat 42.5077043
Long -71.2421914
Select Work Order
Work Order | Assigned | Activity | Comp Type | Address
1111 | Joe Brown | Leak Inspection | Pipeline | 123 Some Street, Newtville, IL
1117 | Unassigned | Keyhole Inspecti | Pipeline | 789 Some Hwy, Newtville, IL
1118 | Joe Brown | First Response | Pipeline | 100 Rural Road, Newtville, IL
Start Report
Cancel
Clear All Filters


Smart Cloud
Component Selection
Set Component Filters
Distance from Current Location
Within 500 ft
Lat 42.5077043
Long -71.2421914
Component Type
Pipeline – Plastic – 4"
Remove
Component ID Mask
7
Select Component
ID | Type | Material | Size (in) | Distance (ft)
72166 | Pipeline | Plastic | 4 | 180
71553 | Pipeline | Plastic | 4 | 320
71486 | Pipeline | Plastic | 4 | 392
Select
Add New Component
Cancel
Clear All Filters


Smart Cloud
Component Data
Component ID 72166
Change
Type Pipe
Physical Address
Address 1 19 Crosby Drive
Address 2
City Bedford
State MA
Zip Code 01730
Geo-Location
Latitude 40.60401754
Longitude -74.10271283
Pipe Characterization
Installation Date 07/29/89
Method Open Trench Drilling
Facility Type Service
Depth 18 in
Description for "Other" facility type
Pipe Material Plastic
Plastic Type AldylA
Diameter 4.0 in
Standard Diameter Ratio 20.0
Max Operating Pressure 50.0 psig
Tee on this segment? Yes No
LDIW in this segment? Yes No

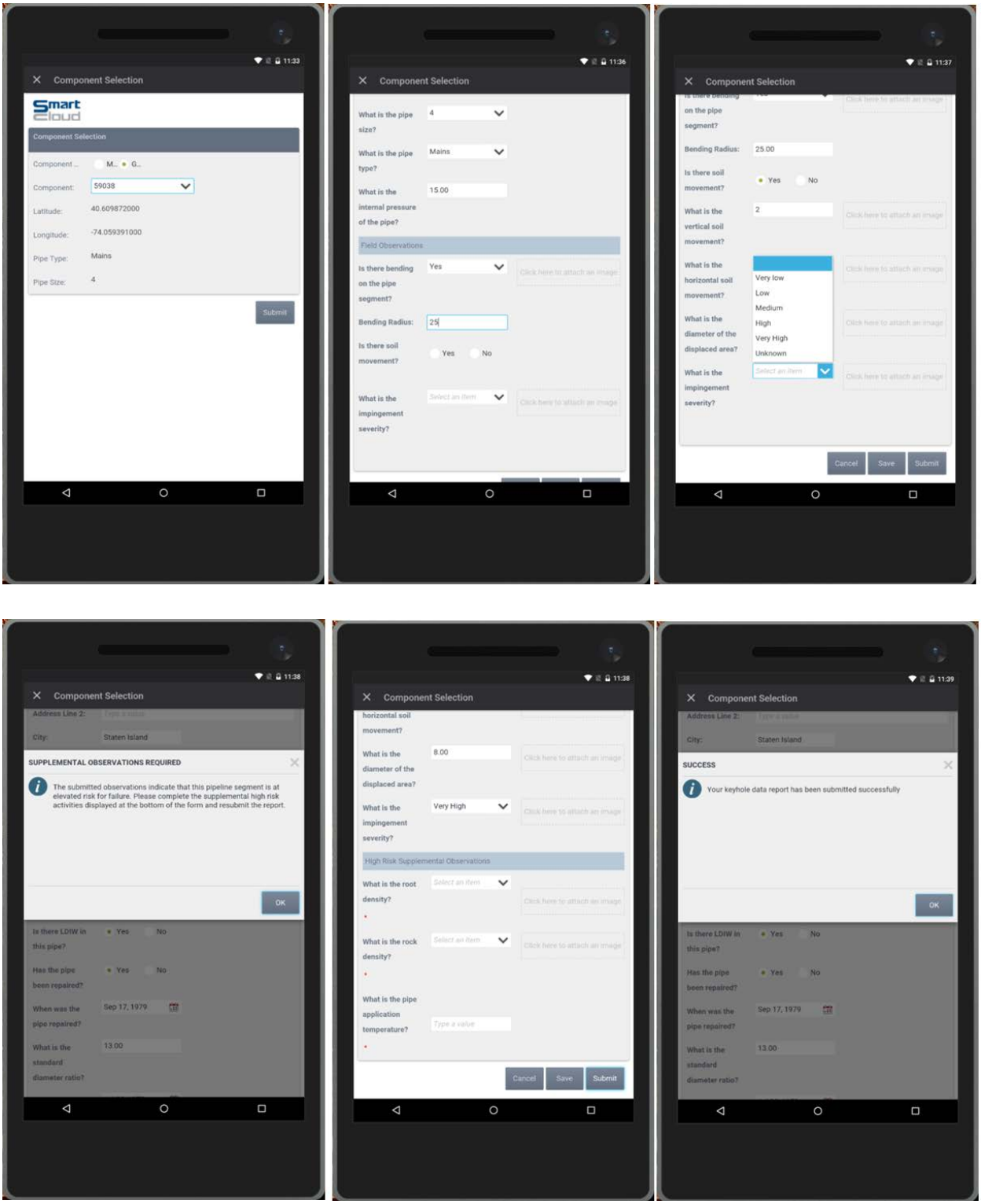


Figure 7-30. Screen captures of Keyhole Inspection Form mobile application

System Integration Testing

In system integration testing, keyhole inspection form was combined with database and Bayesian network in the server and tested as a single system. Table below describes the steps taken during the integration testing and lists the pass/fail grade for each step.

Step #	Step Description	Test	Expected Results	Test Pass/Fail
1	Deploy the latest Keyhole Inspection Form in cloud server	Run unit test for integrity checks	Provides latest version of the application for installation in mobile devices	Pass
2	Create database table to store the keyhole inspection data	Run unit test queries	Stores data collected from Keyhole inspection form	Pass
3	Install the SmartForms application in mobile that includes keyhole leak inspection form	Install/update the application in android phone	Pulls latest version of the application from server and installs it in the mobile device	Pass
4	Open mobile application and enter keyhole inspection data	Enter keyhole inspection data through the mobile application and press save button in the application	Updates database table in the server with new information	Pass

Validation Testing

Validation testing verifies in detail that the keyhole inspection form meets requirement specification and each data collection field works as intended. The table below lists the data fields that were tested and the pass/fail result of the tests.

Screen	Test Field	Entry	Branching	Attachments	Test Pass/Fail
1	Report Provenance Tab				Pass
		Form: Report Provenance			
2	Pipeline Component Tab				Pass
		Form: Pipeline Component			
3	Type of Job on Site				Pass

	(Options: Abandon Services, Cathodic Protection, Valve Maintenance, CI Joint Repair, Test Hole, Other)	Pulldown text		Notes	
3	Number of Core(s)				Pass
		Numeric		Notes	
3	Core Diameter				Pass
	(Options: 18", 24")	Pulldown text		Notes	
3	Core Thickness				Pass
		Numeric, inches		Notes	
3	Vacuum Only				Pass
	(Options: Yes, No)	Button selection		Notes	
3	Depth of Facility				Pass
		Numeric, inches		Notes	
3	Operating Pressure				Pass
		Numeric, psig Not Available		Notes	
3	Maximum Operating Pressure				Pass
		Numeric, psig			
3	Soil Temperature				Pass
	Converted to C for BN	Numeric, °F Not Available		Notes	
3	Pipe Condition				Pass
	(Options: Good, Break, Damage, Light Corrosion, Heavy Corrosion, Thru Corrosion, Joint Flaw, Other)	Pulldown text		Picture, Notes	
3	External Coating				Pass
	(Options: X-tru coat, Coal tar enamel, Tape wrap, FBE, Powder Coating, Mastic, Wax tape, Not Available)	Pulldown text - Any Known Type	Expose: Coating Condition Expose: Adhesion Status	Picture, Notes	
		Pulldown text - Not Available			
3	Coating Condition		Conditional		Pass
	(Options: Good, Bare, Damage, Poor Joint, Deteriorated)	Pulldown text		Picture, Notes	
3	Coating Adhesion		Conditional		Pass
	(Options: Bonded, Disbonded)	Button selection		Notes	
3	External Coating				Pass
	(Options: X-tru coat, Coal tar enamel, Tape wrap, FBE, Powder Coating, Mastic, Wax Tape, Not Available)	Pulldown text		Picture, Notes	
3	Coating Condition				Pass
	(Options: Good, Bare, Damage, Poor Joint, Deteriorated)	Pulldown text		Picture, Notes	
3	Impingement Depth				Pass

	Convert to impingement ratio for classification BN Classifications ([0-0.2],[0.2-0.4],[0.4-0.6],[0.6-0.8],[0.8+]	Numeric, inches		Picture, Notes	
3	Bending on Pipe Segment?				Pass
	(Options: Yes, No, Not Available)	Pulldown - Yes	Expose: Bending Radius	Picture, Notes	
		Pulldown - No			
		Pulldown – Not Available			
3	Bending Radius		Conditional		Pass
		Numeric, feet			
4	Soil Type				Pass
	(Options: Standard, Clay, Gravel, Sand, Peat)	Pulldown text		Picture, Notes	
4	Root Density				Pass
	(Options: Very Low, Low, Medium, High, Very High, Not Available)	Pulldown selection		Picture, Notes	
4	Rock Density				Pass
	(Options: Very Low, Low, Medium, High, Very High, Not Available)	Pulldown selection		Picture, Notes	
4	Soil Movement				Pass
	(Options: Yes, No)	Button selection - Yes	Expose: Vertical Soil Movement Expose: Horizontal Soil Movement Expose: Diameter of Displaced Area	Picture, Notes	
		Button selection - No			
4	Vertical Soil Movement		Conditional		Pass
		Numeric, feet		Picture, Notes	
4	Horizontal Soil Movement		Conditional		Pass
		Numeric, feet		Picture, Notes	
4	Diameter of Displaced Soil Area		Conditional		Pass
		Numeric, feet		Picture, Notes	
4	Prior Leak Repair?				Pass
	(Leak Repair Report ID, if found "No Repair Recorded", if no prior report "Not Available" if no CRex connectivity)	Report ID	Expose: Repair Date		
		"No Repair Recorded"			
		"Not Available"			
4	Repair Date		Conditional		Pass
		Date (YYMMDD)			
5	Leak Found?				Pass

	(Options: Yes, No)	Button selection	Expose: Leak Grade Expose: Leak Size Characterization	Notes	
5	Equipment Used				Pass
	(Options: Flame Ionization, Remote Methane Leak Detection, Other)	Pulldown text - Flame Ionization		Notes	
		Pulldown text - Remote Methane Leak Detection			
		Pulldown text - Other	Expose: Other Equip Type Description		
5	Wind Speed				Pass
		Numeric, m/s Not Available		Notes	
5	Wind Direction				Pass
	(Options: N, NE, E, SE, S, SW, W, NW, Not Available)	Pulldown text		Notes	
5	Leak Grade		Conditional		Pass
	(Options: Grade 1, Grade 2, Grade 3)	Pulldown text		Notes	
5	Leak Size Characterization		Conditional		Pass
	(Options: Diameter, Length)	Button selection - Diameter	Expose: Leak Size - Diameter	Notes	
		Button selection - Length	Expose: Leak Size - Length		
5	Leak Size - Diameter		Conditional		Pass
		Numeric, in		Picture, Notes	
5	Leak Size - Length		Conditional		Pass
		Numeric, in		Picture, Notes	

Risk Calculation Testing

To demonstrate how keyhole inspection form helps to refine the risk profiles of the gas distribution network with latest data, we selected five pipeline segments randomly and calculated mean life expectancy and risk score before and after the new data was collected with keyhole inspection form. **Table 7-4** and **Table 7-5** show the data before and after the data is collected. Then, the original and updated data are given as input for the developed Bayesian network, which estimates the life expectancy for each pipe segment. The mean life expectancy is the direct measure for the pipe failure likelihood, while the consequence of the failure is determined by the type of buildings or facilities. Given the likelihood and consequence, the risk is simply the product of those two measures. The computed risk for the original and update data is shown in **Table 7-6**.

Table 7-4. Data before Keyhole inspection collects new information

ID	Years after Repair	Root	Rock	Horiz Soil_ Mov	Vert_Soil _Mov	Diameter of Displaced Area	Pipe Size	Pressure	If LDW	If Repaired	Leak Type
1	30	Med	Med	8	11	25	<= 2 inch	45	True	True	Type 3
2	31	Med	Very High	0	5	8	<= 2 inch	45	True	True	Type 3
3	26	Med	Low	8	7	8	<= 1 inch	45	True	True	Type 3
4	32	Med	Very Low	1	6	2	<= 1 inch	55	True	True	Type 3
5	31	Med	Med	3	6	25	<= 4 inch	55	True	True	Type 3

Table 7-5. Latest data from Keyhole inspection

ID	Years after Repair	Root	Rock	Horiz Soil_ Mov	Vert_ Soil_ Mov	Diameter of Displaced Area	Pipe Size	Pressure	If LDW	If Repaired	Leak Type
1	33	Med	Med	9	11	26	<= 2 inch	45	True	True	Type 3
2	34	Med	Very High	1	5	8.7	<= 2 inch	45	True	True	Type 1
3	29	Med	High	8	7	8	<= 1 inch	45	True	True	Type 1
4	35	Med	Low	1	9	2.2	<= 1 inch	55	True	True	Type 2
5	34	Med	Med	3	9	27.5	<= 4 inch	55	True	True	Type 2

Table 7-6. Mean life prediction of pipeline segments before and after data is collected with keyhole inspection form

ID	Mean life prediction before data is collected	Mean life prediction after data is collected
1	0.615875	0.740329
2	0.799914	0.799998
3	0.282651	0.357131
4	0.799711	0.79999
5	0.799999	0.8

From **Table 7-7**, the mean life expectancy is decreasing when the data are updated three years later. However, the change in life expectancy is relatively small compared with the original values. Therefore, the updated risks for those 3 segments are not affected by the

newly gathered data. Practically, effective prevention or mitigation measures should be taken if the risks are higher than a threshold value (i.e. risk tolerance).

Table 7-7. Risk after a number of years

	ID	1 yr	2 yrs	3 yrs	4 yrs	5 yrs	10 yrs	15 yrs	20 yrs	30 yrs
Before data is collected with Keyhole Inspection Form	1	3.22E-06	0.000445	0.004234	0.015893	0.038045	0.256674	0.475484	0.615875	0.740329
	2	0.002101	0.078987	0.272732	0.467648	0.605736	0.788866	0.799168	0.799914	0.799998
	3	6.12E-08	1.04E-05	0.000129	0.000621	0.001868	0.029596	0.09568	0.182651	0.357131
	4	0.002505	0.074884	0.249289	0.430559	0.567443	0.780286	0.797916	0.799711	0.79999
	5	0.026625	0.296322	0.566769	0.703688	0.761005	0.799371	0.799979	0.799999	0.8
After data is collected with Keyhole Inspection Form	1	4.43E-06	0.000611	0.005669	0.02074248	0.048468	0.298187	0.522008	0.652887	0.75768
	2	0.003519	0.100391	0.309649	0.5015927	0.629989	0.790707	0.799302	0.799927	0.799998
	3	1.08E-07	1.6E-05	0.000186	0.00085565	0.002484	0.035684	0.109744	0.202996	0.382059
	4	0.001216	0.039984	0.152908	0.29763093	0.431237	0.737363	0.788531	0.797506	0.799818
	5	0.037506	0.345769	0.60884	0.72716615	0.772429	0.799645	0.79999	0.799999	0.8

Testing of First Response Form

Functional Testing

Each functionality of the First Response Form was tested to verify that it meets the requirement specification. An appropriate input is entered in the system and the output is verified by comparing the actual result with the expected result. The table below illustrates the steps taken for functional testing and lists the pass/fail grade for each testing step.

Step #	Step Description	Test Data	Expected Results	Test Pass/Fail
1	Enter login information	username: *admin* password: *Admin12248*	Show first response form main page	Pass
2.a	Filter existing pipelines by barcode	ASTM F2897 barcode	List pipelines matching information in barcode	Pass
2.b	Filter existing pipelines by distance	100 feet	List pipelines within 100 feet of current location	Pass
2.c	Filter existing pipelines by type	Aldyl-A plastic pipe	List all Aldyl-A pipelines	Pass
3.a	Select existing pipeline from filtered list of pipelines	Pipeline with ID '72166'	Show next screen to view historical data of the pipeline selected	Pass
3.b	Add new pipeline information	Pipeline with ID '77667'	Save new pipeline information and show screen to enter data	Pass
4	View historical pipeline data		Show past data and allow updates	Pass
5	Enter new first response data	**Please refer to:**	Collects all first response data. As a next step, enable Save button	Pass
6	Save the entered first response data	Press Save button	Save the data in cloud database	Pass

System Integration Testing

In the system integration testing, First Response Form was combined with database and Bayesian network in the server and tested as a single system. Table below describes the steps taken during the integration testing and lists the pass/fail grade for each step.

Step #	Step Description	Test	Expected Results	Test Pass/Fail
1	Deploy the latest First Response Form in cloud server	Run unit test for integrity checks	Provides latest version of the application for installation in mobile devices	Pass
2	Create database table to store the first response data	Run unit test queries	Stores data collected from first response form	Pass
3	Install the SmartForms application in mobile that includes first response form	Install/update the application in android phone	Pulls latest version of the application from server and installs it in the mobile device	Pass
4	Open mobile application and enter first response data	Enter first response data through the mobile application and press save button in the application	Updates database table in the server with new information	Pass

Validation Testing

Validation testing verifies in detail that the First Response Form meets requirement specification and each data collection field works as intended. The table below lists the data

fields that were tested and the pass/fail result of the tests. First response form is divided into different tabs in android application for ease of usage.

Tab: First Response Header

Screen	Field	Entry	Branching	Attachments	Test Pass/Fail
Hdr	First Response No.				
		FSTR-000000			Pass
Hdr	Work Order No.				
		Text, from WO selection			Pass
Hdr	Component ID				
		Text, from Component tab			Pass

Tab: First Response Footer

Screen	Field	Entry	Branching	Attachments	Test Pass/Fail
Hdr	First Response No.				
		FSTR-000000			Pass
Hdr	Work Order No.				
		Text, from WO selection			Pass
Hdr	Component ID				
		Text, from Component tab			Pass

Tab: First Response

Screen	Field	Entry	Branching	Attachments	Test Pass/Fail
1	Report Provenance Tab				
		Form: Report Provenance			Pass
2	Component Tab				
		Form: Basic Info (top half) Based on Component Type (bottom half): Form: Pipeline, Ind Meter, or Res Meter	Based on type, Expose: Pipeline, Industrial Meter, or Residential Meter tab		Pass
3	Report Type				
	(Options: Leak Response, Exposed Metal Pipe)	Button selection - Leak Response	Expose: Leak Response tab		Pass
		Button selection - Exposed Metal Pipe	Expose: Exposed Pipe tab		
4a	Leak Response Tab		Conditional		
		Form: Leak Response			Pass
4b	Exposed Pipe Tab		Conditional		
		Form: Exposed Pipe			Pass

Tab: Leak Response

Screen	Field	Entry	Branching	Attachments	Test Pass/Fail
1	Leak Grade				
	(Options: Grade 1, Grade 2, Grade 3)	Pulldown text		Notes	Pass
1	Wind Speed	Not Available			
		Numeric, m/s		Notes	Pass
1	Wind Direction				
	(Options: N, NE, E, SE, S, SW, W, NW, Not Available)	Pulldown text		Notes	Pass
1	Ambient Temperature	Not Available			
	Converted to C for Bayesian Network	Numeric, °F		Notes	Pass
1	Equipment ID				
		Text		Notes	Pass
1	Inside Investigation Needed				
	(Options: Yes, No)	Button selection - Yes	Expose: Inside Investigation Performed	Notes	Pass
		Button selection - No			
1	Inside Investigation Performed		Conditional		
	(Options: Yes, No, No Access)	Button selection		Picture, Notes	Pass
1	Soap Test				
	(Options: Yes, No)	Button selection - Yes	Expose: Test Type Expose: Start Time Expose: End Time Expose: Pressure Expose: Results Expose: Length of Pipe	Notes	Pass
		Button selection - No			
1	Test Type		Conditional		
	(Options: Air, Gas, Water, Nitrogen)	Pulldown text		Picture, Notes	Pass
1	Start Time		Conditional		
		Timestamp			Pass
1	End Time		Conditional		
		Timestamp			Pass
1	Pressure		Conditional		
		Numeric, psig Not Applicable		Notes	Pass
1	Test Result		Conditional		
	(Options: Pass, Fail)	Button selection		Notes	Pass
1	Length of Pipe		Conditional		
		Numeric, ft Not applicable		Picture, Notes	Pass

Tab: Exposed Pipe

Screen	*Field*	*Answers*	*Branching*	*Attachments*	*Test Pass/Fail*
1	Length of Pipe Inspected				
		Numeric, ft			Pass
1	Exposed Pipe Accessibility				
	(Options: Accessible, Inaccessible)	Button selection		Picture, Notes	Pass
1	First PSP Reading				
		Numeric, volts			Pass
1	First PSP Read Latitude				
		Numeric Use Current Location			Pass
1	First PSP Read Longitude				
		Numeric Use Current Location			Pass
1	First PSP Outside Limit?				
	(Options: Outside Limits, Within Limits)	Text - "Outside Limits"	Expose: Tech Called Expose: Tech Callout Time		Pass
		Text- "Within Limits"			
1	Tech Called (per GOS 2575.2800)		Conditional		
		Text			Pass
1	Tech Callout Time		Conditional		
		Timestamp			Pass
2	External Coating				
	(Options: X-tru coat, Coal tar enamel, Tape wrap, FBE, Powder Coating, Mastic, Wax tape, Not Available)	Pulldown text - Any Known Type	Expose: Coating Condition Expose: Adhesion Status	Picture, Notes	Pass
		Pulldown text - Not Available			
2	Coating Condition		Conditional		
	(Options: Good, Bare, Damage, Poor Joint, Deteriorated)	Pulldown text		Picture, Notes	Pass
2	Coating Adhesion		Conditional		
	(Options: Bonded, Disbonded)	Button selection		Notes	Pass
2	Pipe Damage				
	(Options: Damaged, Not Damaged)	Button selection - Damaged	Expose: Damage Requires Remediation Expose: Describe Damage	Picture, Notes	Pass
		Button selection - Not Damaged			
2	Damage Requires Remediation? (per GOS 2575.1700)		Conditional		
	(Options: Yes, No)	Button selection		Notes	Pass
2	Describe Pipe Damage		Conditional		
		Text			Pass

3	External Corrosion?				
	(Options: Corrosion, No Corrosion)	Button selection - Corrosion	Expose: External Corrosion Type Expose: Corrosion Requires Remediation	Picture, Notes	Pass
		Button selection - No Corrosion			
3	External Corrosion Type		Conditional		
	(Options: Surface Rust, Isolated Pit, Multiple Pits, General Corrosion, Not Available)	Pulldown text - Isolated Pit - Multiple Pits	Expose: Maximum Pit Depth Expose: Pit Circumference Expose: Pit Position	Picture, Notes	Pass
		Pulldown text - Multiple Pits	Expose: Distance between Pits Expose: Length of Pitting		
		Pulldown text - all others			
3	Maximum Pit Depth		Conditional		
		Numeric, in		Picture, Notes	Pass
3	Pit Circumference		Conditional		
		Numeric, in		Picture, Notes	Pass
3	Distance between Pits		Conditional		
		Numeric, ft		Picture, Notes	Pass
3	Length of Pitting on Pipe		Conditional		
		Numeric, ft		Picture, Notes	Pass
3	Position of Pitting		Conditional		
		Text box		Picture	Pass
3	Corrosion Requires Remediation? (per GOS 2600.1900)		Conditional		
	(Options: Yes, No)	Button selection		Notes	Pass
4	Pipe Interior Visible?				
	(Options: Visible, Not Visible	Button selection - Visible	Expose: Retired Pipe Expose: Coupon Expose: Internal Corrosion	Picture, Notes	Pass
		Button selection - Not Visible			
4	Retired Pipe		Conditional		
	(Options: Yes, No)	Button selection		Notes	Pass
4	Coupon		Conditional		
	(Options: Yes, No)	Button selection - Yes	Expose: Coupon ID	Picture, Notes	Pass
		Button selection - No			
4	Coupon ID		Conditional		
		Text			Pass
4	Internal Corrosion?				
	(Options: Corrosion, No Corrosion, Not Available)	Button selection - Corrosion	Expose: Internal Corrosion Type	Picture, Notes	Pass

			Expose: Corrosion Requires Remediation		
		Button selection - No Corrosion			
		Button selection - Not Available	Expose: Not Available Explanation		
4	Not Available Explanation		Conditional		
		Text box			Pass
4	Internal Corrosion Type		Conditional		
	(Options: Surface Rust, Isolated Pit, Multiple Pits, General Corrosion)	Pulldown text - Isolated Pit - Multiple Pits	Expose: Maximum Pit Depth Expose: Pit Circumference Expose: Pit Position	Picture, Notes	Pass
		Pulldown text - Multiple Pits	Expose: Distance between Pits Expose: Length of Pitting		
		Pulldown text - all others			
4	Maximum Pit Depth		Conditional		
		Numeric, in		Picture, Notes	Pass
4	Pit Circumference		Conditional		
		Numeric, in		Picture, Notes	Pass
4	Distance between Pits		Conditional		
		Numeric, ft		Picture, Notes	Pass
4	Length of Pitting on Pipe		Conditional		
		Numeric, ft		Picture, Notes	Pass
4	Position of Pitting		Conditional		
		Text box		Picture	Pass
4	Corrosion Requires Remediation? (per GOS 2600.1900)		Conditional		
	(Options: Yes, No)	Button selection		Notes	Pass
5	Cast Iron Graphitization?		Conditional Tab: Pipe Material: Cast Iron		
	(Options: Yes, No)	Button selection - Yes	Expose: Graphitization Location Expose: Extent of Graphitization	Picture, Notes	Pass
		Button selection - No			
5	Graphitization Location		Conditional		
		Text			Pass
5	Extent of Graphitization		Conditional		
		Text			Pass
6	Second PSP Reading				
		Numeric, volts			Pass
6	Second PSP Read Latitude				
		Numeric Use Current Location			Pass

6	Second PSP Read Longitude				
		Numeric Use Current Location			Pass
6	Second PSP Outside Limit?				
	(Options: Outside Limits, Within Limits)	Text - Outside Limits	Expose: Tech Called Expose: Tech Callout Time		Pass
		Text- Within Limits			
6	Tech Called (per GOS 2575.2800)		Conditional		
		Text			Pass
6	Tech Callout Time		Conditional		
		Timestamp			Pass

Testing of Audit Report Form

Functional Testing

Each functionality of the Audit Report Form was tested to verify that it meets the requirement specification. An appropriate input is entered in the system and the output is verified by comparing the actual result with the expected result. The table below illustrates the steps taken for functional testing and lists the pass/fail grade for each testing step.

Step #	Step Description	Test Data	Expected Results	Test Pass/Fail
1	Enter login information	username: *admin* password: *Admin12248*	Show audit report main page	Pass
2	View annual report data		Show gas distribution system data from past year	Pass
3	Enter data for any empty fields	**Please refer to:**	**Collects.** As a next step, enable Save button	Pass
4	Save the completed audit report form	Press Save button	Save the data in cloud database	Pass

Figure 7-31 shows the screenshot of Annual Report Form design.

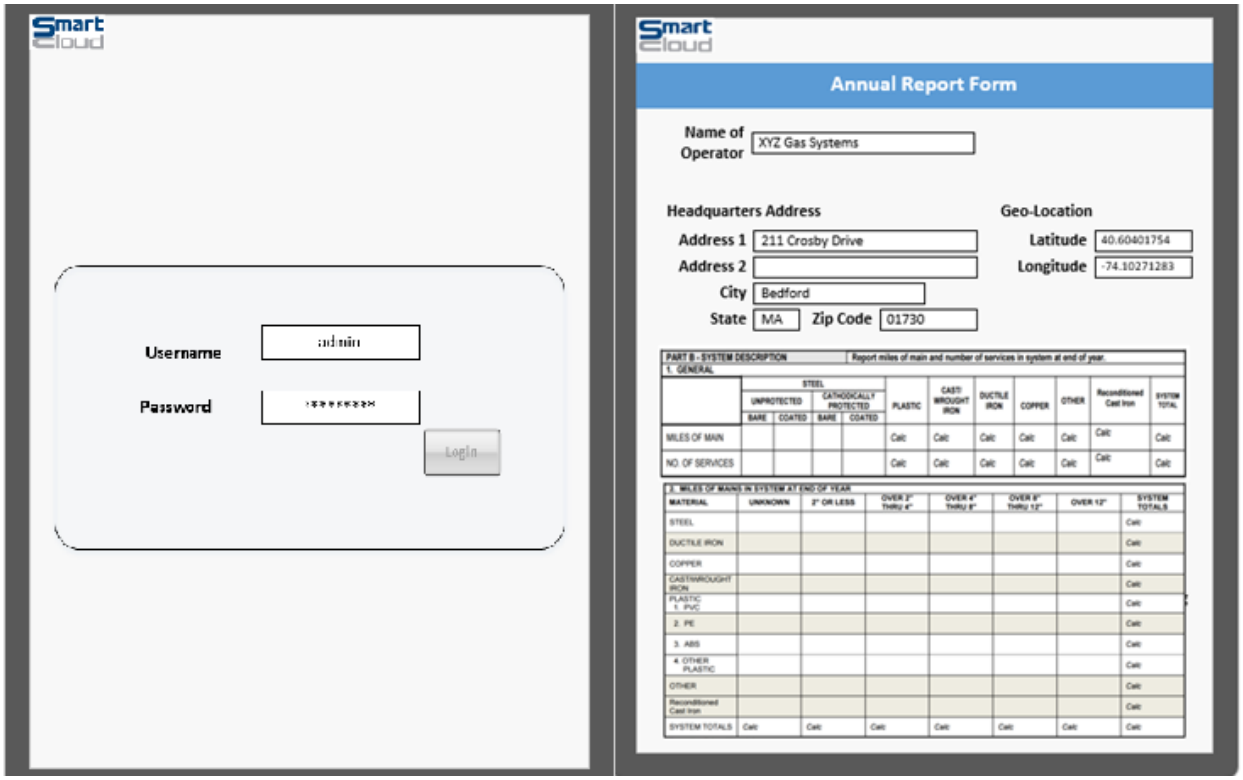


Figure 7-31. Annual Report Form login page and main section

System Integration Testing

In the system integration testing, Audit Report Form was connected to the database in the server and tested as a single system. Table below describes the steps taken during the integration testing and lists the pass/fail grade for each step.

Step #	Step Description	Test	Expected Results	Test Pass/Fail
1	Deploy the latest annual report form in cloud server	Run unit test for integrity checks	Provides latest version of the application for report generation	Pass
2	Create database table to store the annual report data	Run unit test queries	Stores data derived from the main database of gas distribution system	Pass
3	Open and complete Audit report form	Verify data displayed in the audit form, complete any remaining empty fields, and press save button in the application	Updates database table in the server with new information	Pass

Validation Testing

Validation testing verifies in detail that the Audit Report Form meets requirement specification and each data collection field works as intended. The table below lists the data fields that were tested and the pass/fail result of the tests.

Audit for Integrity Management Program (Process, Inspection, Mitigation, and Prevention) as prescribed in ASME B31.8S standard

Integrity Management Program Area	Audit Questions	Test Pass/Fail
Pipe Segmentation	Has the whole pipeline system been segmented? • How many miles of pipeline were investigated? • Was it different from the value specified by the requirements? What were the company specific integrity management goals/objectives? Did the operator consider those goals when applying the process? Was there a new pipeline becoming a part of an integrity management program? • If so, were the functional requirements, including prevention, detection, and mitigation considered? Is the risk assessment results applicable to the pipe within the segment? • If not, what was the reason for limiting its application?	Pass
Threat Identification	Did the operator identify/evaluate all the threats? • If not, which one(s) of following was the reason for excluding the threat/threats? • There is no history of a threat impacting the particular segment or pipeline system • The threat is not supported by applicable industry data or experience • The threat is not implied by related data elements • The threat is not supported by like/similar analyses • The threat is not applicable to system or segment operating conditions Were there multiple threats occurring on the pipe segment? • If so, was the interaction considered?	Pass
Data Collection and Integration	Did the operator have a comprehensive plan for collecting, reviewing, and analyzing all data sets? Did the operator include all possible data sources specified in ASME B31.8S? • If not, what was the reason for excluding the specific data source? Did the operator check the quality and deficiency of the collected data?	Pass

	• Did the operator follow the ASCE, PHMSA ADB, ASTM GARP, and ISO standards when checking the quality? • Did the operator follow the prescriptive process when the data quality and sufficiency was below requirement?	
Risk Assessment	Did the operator consider all the consequence factors specified in ASME B31.8S when doing risk assessment? • If not, what was the reason for excluding the specific factor? Did the risk assessment possess the following characteristics? a. Attributes(Defined logic and be structured) b. Resources(Adequate personnel and time) c. Operating/mitigation history(Consider the frequency and consequence of past events) d. Predictive capability(Identify threats not considered) e. Risk confidence (All data verified and checked) f. Feedback(Continuous improvement) g. Documentation(Thoroughly documented) h. What if (Structure necessary to perform "what if" calculation) i. Weighting factors (All threats and consequences weighted) j. Structure (Compare, rank risk results to prioritize integrity management decision making) k. Segmentation(Incorporate sufficient resolution of pipe pipeline segment size) Was the result from risk assessment consistent with the industry's experience? Was the result from risk assessment used to prioritize initial integrity assessment?	Pass
Integrity Assessment	Did the operator follow the standard ASME B31.8S for the selecting of integrity inspection tool for each threat type? Were the techniques used not published by consensus standard? • If so, did the operator follow the performance requirements of ASME B31.8S code and confirm the validity of this approach? Were indications of defects discovered during inspection?	Pass

	• If so, was the indication examined and evaluated to determine if they are actually defects or not? Was any abnormality discovered using the inspection method? • What was the number of anomalies found during inspection? • Were these anomalies addressed properly? Was the information updated periodically? • If so, what is the frequency/interval for the information updating? Was the plan flexible enough to incorporate new information acquired during integrity management? Was the risk reassessment conducted periodically? • If so, what is the frequency/interval for the reassessment? Were the results from risk and integrity assessment considered when prioritizing inspections and mitigation activities?	
Response to Integrity Assessment	Was the repair activity made in accordance with ASME B31.8S? Did the operator evaluate prevention techniques that prevent future deterioration of the pipeline? Was the information obtained from risk and integrity assessment, mitigation activities added to database? Was the plan or program updated when additional information was acquired and incorporated?	Pass

Performance Measures	Was enough time given to collect the data for measures selection?
	Did the selected performance measures possess the general characteristics, including simple, measurable, attainable, relevant, permit timely evaluation?
	• If not, which one(s) the following characteristics was missing?
	  o Simple
	  o Measurable
	  o Attainable
	  o Relevant
	  o Permit timely evaluation
	Did the operator evaluate the performance measures periodically?
		Pass

	• If yes, what is the frequency/interval of the evaluation? (At least annually) Did the operator conduct internal benchmarking and audit conducted on the selected measures? Were the results from performance measures utilized to improve the integrity management process?	
Change of Plan	Were there operational changes for significant pressure, pressure cycles, and pressure fluctuation on the pipe segment? • If yes, did the operator consider the fatigue failure mechanism? Have the changes on technology, physical, process, organization been addressed? Have the changes been incorporated into future risk assessment? Has the plan change been monitored to ensure the performance measures effective? Was benchmarking with other gas pipeline operators used? • If so, was the performance measure derived from it carefully evaluated to ensure the comparisons are valid? Were recommendations for changes based on the analysis of performance measures and audits?	Pass
Communication Plan	Did the operator keep the public, emergency responder, location official, jurisdictional authorities, and stakeholder informed during the process? Did the communications take place periodically? Was there a written procedure for the change of plan? Were the system changes listed below reflected in the integrity management program? a. Change in land use b. Changes to the MAOP c. Daily changes in operating pressure Were changes identified and reviewed before implementation? Did the personnel understand the whole plan change procedures? Did the operator investigate the effectiveness of the plan? Did the operator refresh the training for personnel after the change of plan?	Pass

	Were the appropriate stakeholders given the opportunity to participate the risk assessment process? Have the operator documented the changes, new technologies, and public awareness?	
Quality Control Plan	Were there quality control activities listed below in the integrity management plan? a. Identify the processes that will be included in the quality program b. Determine the sequence and interaction of these process c. Determine the criteria and methods needed to ensure that both the operation and control of these processes are effective d. Provide the resources and information necessary to support the operation and monitoring of these processes e. Monitor, measure, and analyze these processes f. implement actions necessary to achieve planned results and continued improvement of these processes Were the relevant documents controlled and maintained at appropriate locations for the duration of the program? Were the responsibilities and authorities clearly, formally defined? Were results of the quality control plan reviewed at predetermined intervals? Were the personnel involved in the program competent and qualified to execute the activities within the program? Did the operator monitor the program to show it was being implemented according to plan? Were control points, criteria, and/or performance metrics defined? Were corrective actions to improve quality plan documented and their implementation monitored? Were outside resources used to conduct any process that affected the quality of the program? • If so, did the operator ensure control of such processes and document them within the quality program?	Pass

Query Testing

In this test, queries developed to fill PHMSA annual audit report for gas distribution pipeline titled "PHMSA F 7100.1-1 (Annual report form)" were tested. Sections A, B and C of the form was automatically filled with the aggregated data from the database.

Conclusions

The tasks outlined in this section show that it is feasible to integrate the risk assessment tools developed in this project into a commercially available artificial intelligence platform designed to provide enterprise level decision support. The tasks discussed above addressed data flows and the provision of intelligent feedback to the operators given the data inputs and the analyses run on the inputs. We have demonstrated the system components necessary to arrive at system wide, and individual segment fitness for service information needed to drive rate cases, replacement programs and asset maintenance programs.

8. Feasibility of Using a Short-Term Pressure Test to Determine Fitness for Service in the Context of Slow Crack Growth

Kiefner and Maxey, 2000, published a report, "The Benefits and Limitations of Hydrostatic Testing", that gives a good overview of the motivation behind using a short-term pressure test to determine fitness for service of steel pipelines [89]. Some select portions of this report are presented in the inserts and **Figure 8-1** below:

Both field experience and full-scale laboratory tests have revealed much about the benefits and limitations of hydrostatic testing. Among the things learned were the following:

- Longitudinally oriented defects in pipe materials have unique failure pressure levels that are predictable on the basis of the axial lengths and maximum depths of the defects and the geometry of the pipe and its material properties[(2)].
- The higher the test pressure, the smaller will be the defects, if any, that survive the test.
- With increasing pressure, defects in a typical line-pipe material begin to grow by ductile tearing prior to failure. If the defect is close enough to failure, the ductile tearing that occurs prior to failure will continue even if pressurization is stopped and the pressure is held constant. The damage created by this tearing when the defect is about ready to fail can be severe enough that if pressurization is stopped and the pressure is released, the defect may fail upon a second or subsequent pressurization at a pressure level below the level reached on the first pressurization. This phenomenon is referred to as a pressure reversal[(3, 4)].
- Testing a pipeline to its actual yield strength can cause some pipe to expand plastically, but the number of pipes affected and the amount of expansion will be small if a pressure-volume plot is made during testing and the test is terminated with an acceptably small offset volume or reduction in the pressure-volume slope[(5)].

TEST-PRESSURE-TO-OPERATING-PRESSURE RATIO

The hypothesis that "the higher the test-pressure-to-operating-pressure ratio, the more effective the test", is validated by Figure 1. Figure 1 presents a set of failure-pressure-versus-defect-size relationships for a specific diameter, wall thickness, and grade of pipe. A great deal of testing of line-pipe materials over the years has validated these curves[2]. Each curve represents a flaw with a uniform depth-to-wall-thickness ratio. Nine such curves are given (d/t ranging from 0.1 to 0.9).

Consider the maximum operating pressure for the pipeline (the pressure level corresponding to 72 percent of SMYS). That pressure level is represented in Figure 1 by the horizontal line labeled MOP. At the MOP, no defect longer than 10 inches and deeper than 50 percent of the wall thickness can exist. Any such defect would have failed in service. Similarly, no defect longer than 4 inches and deeper than 70 percent of the wall thickness can exist, nor can one that is longer than 16 inches and deeper than 40 percent of the wall thickness.

By raising the pressure level above the MOP in a hydrostatic test, the pipeline's operator can assure the absence of defects smaller than those that would fail at the MOP. For example, at a test pressure level equivalent to 90 percent of SMYS, the largest surviving defects are determined in Figure 1 by the horizontal line labeled 90 percent of SMYS. At that level, the longest surviving defect that is 50 percent through the wall can be only about 4.5 inches. Compare that length to the length of the longest possible 50-percent-through flaw at the MOP; it was 10 inches. Alternatively, consider the minimum survivable depth at 90 percent SMYS for a 10-inch-long defect (the size that fails at the MOP if it is 50 percent through the wall). The survivable depth is only about 32 percent through the wall. By a similar process of reasoning, one can show that even smaller flaws are assured by tests to 100 or 110 percent of SMYS (the horizontal lines drawn at those pressure levels on Figure 1).

The point is that the higher the test pressure (above MOP), the smaller will be the possible surviving flaws. This fact means a larger size margin between flaw sizes left after the test and the sizes of flaws that would cause a failure at the MOP. If surviving flaws can be extended by operating pressure cycles, the higher test pressure will assure that it takes a longer time for these smaller flaws to grow to a size that will fail at the MOP. Thus, Figure 1 provides proof of the validity of the hypothesis (i.e., the higher the test-pressure-to-operating-pressure ratio, the more effective the test).

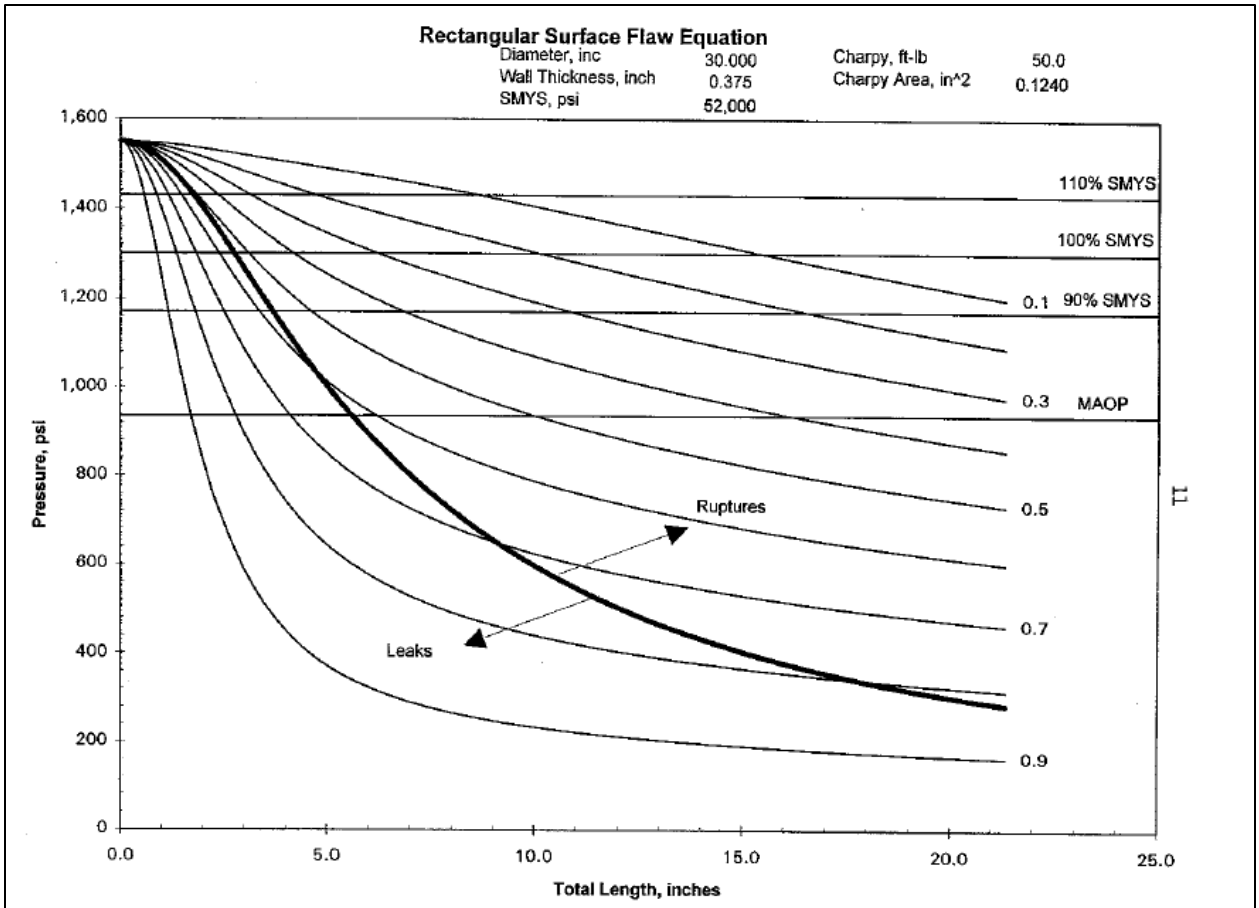


Figure 1. Impact of Test Pressure Levels on Margin of Safety

Figure 8-1. Kiefner [89] Figure 1.

The feasibility of the approach is clearly based on how defects propagate in typical pipeline steels, i.e. cracks will propagate a small distance into the pipe wall under load, and arrest (remain stable) until the rupture boundary for the pipeline steel is reached. In **Figure 8-1** Kiefner and Maxey show an illustrative example for a steel with known yield strength and toughness. There is a well understood relationship between the crack length, the ratio of crack depth to wall thickness and the rupture boundary for each combination of pipe diameter, wall thickness, yield strength and toughness that can be used to calculate the leak rupture boundary for steel pipe[11].

---

11 http://www.ttoolboxes.com/documents/OTD_TOC/13-0002_4_TOC.pdf
Leak Rupture Boundary Calculator and Training Manual, and Final Report
OTD Project Number: 4.9.a
Document Number: OTD-13/0002 and OTD-13/0004
Daniel Ersoy, Ernest Lever

We will now investigate whether a similar approach is feasible for vintage polyethylene pipelines.

Understanding Highly Localized Ductile Creep Failure (SCG)

GTI has worked on several projects designed to test the impact of different fluid media on the ductile and SCG failure modes of various polyethylene materials. To accomplish this in an effective and consistent manner a unique Universal Test Vessel (UTV) was designed and specialized test rigs were built to carefully data log the tests. The basic UTV design is shown in **Figure 8-2**.

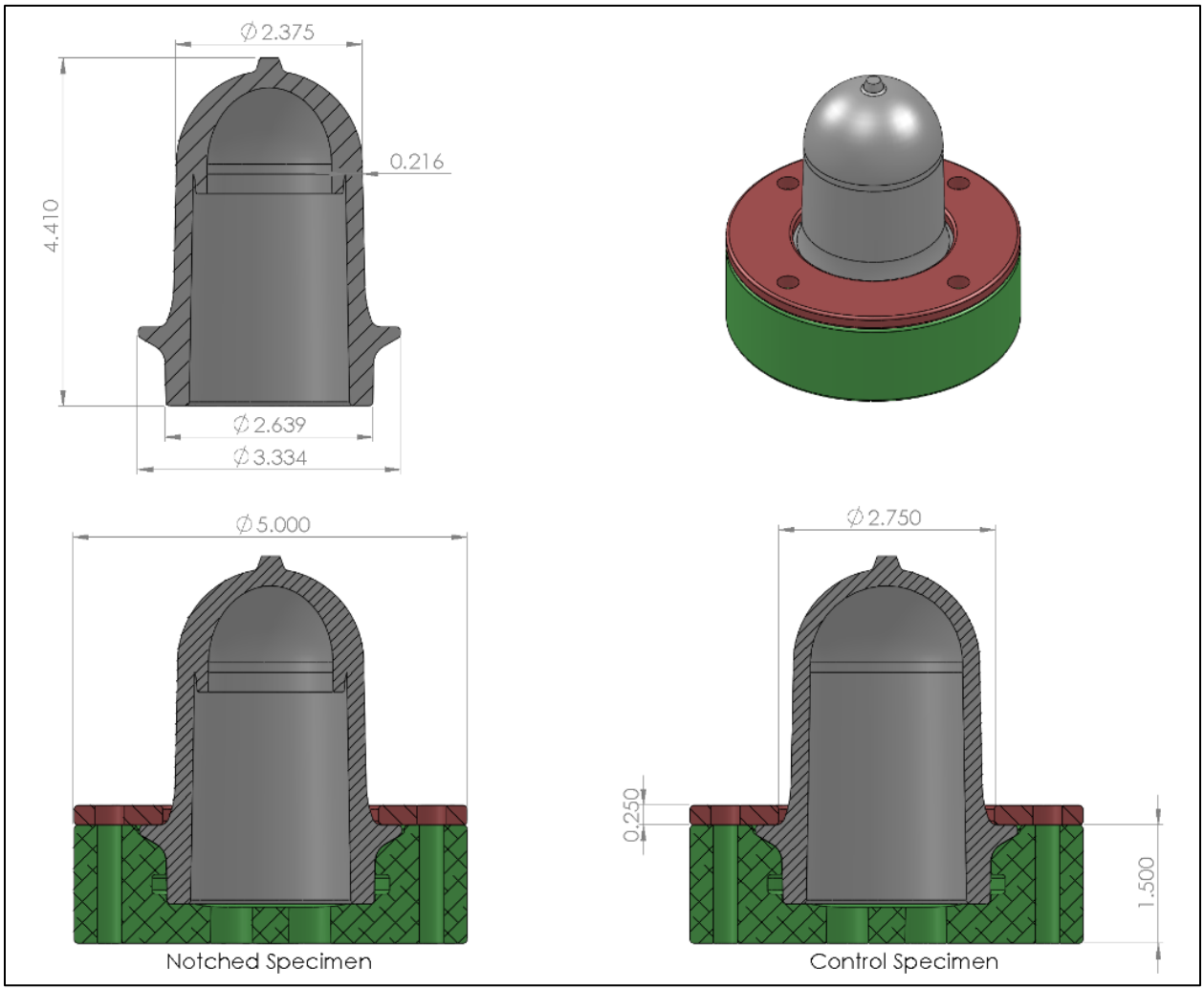


Figure 8-2. Design of Universal Test Vessell (UTV), with and without SIF notch

The UTV has two configurations:

1. *A hemispherical domed version designed to replicate the hoop stress condition in pipe*

2. *A version with a thickened dome and razor sharp molded notch designed to introduce a severe SIF to encourage SCG.*

Specimens can be molded from pellets, or reground pipe exhumed from service. Both methods were used in different projects with good success.
Ductile failure in a UTV with no stress riser notch is shown in **Figure 8-3**, and ductile failure from a stress riser notch is shown in **Figure 8-4.**

Figure 8-3. Ductile failure in un-notched UTV

The ductile failure originating at the stress riser notch shown in **Figure 8-4** is similar to the ductile and mixed-mode failures seen in the CPF testing shown cyclic pressure testing of notched pipe specimens

Figure 8-4. Ductile failure initiated at the stress riser notch

True SCG failures were induced in notched UTV molded from first generation MDPE reground from pipe that had been in service for 40 years. The sensitivity of the test equipment and the leak detection approach used allowed very early identification of wall breach. Wall breaches in the range of 200 – 300 microns were routinely detected as shown in **Figure 8-5.**

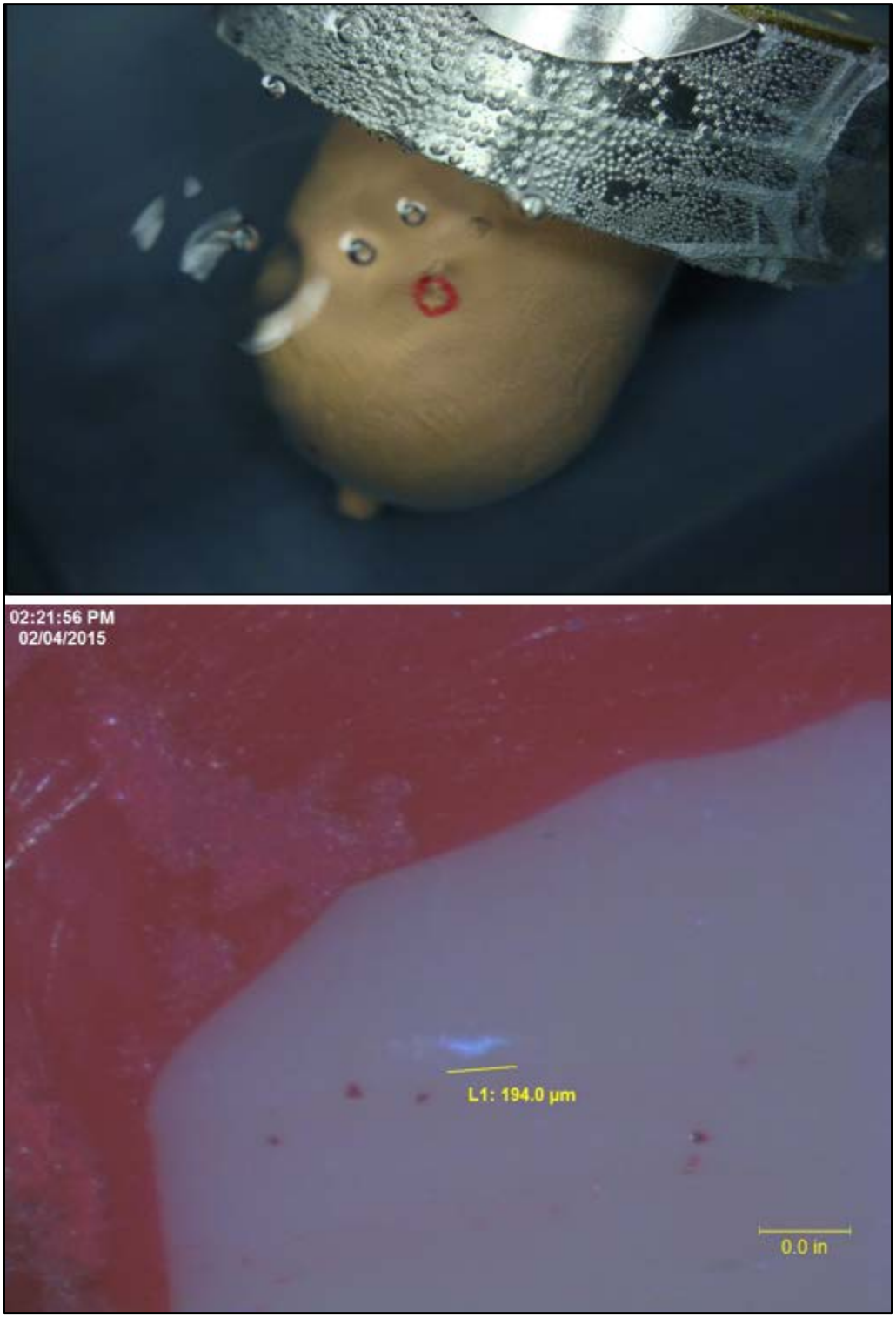


Figure 8-5. Early detection of wall breach due to SCG

Detailed FEM analysis of the stress regime around the notch tip was carried out using a proprietary and highly detailed constitutive model for polyethylene materials. The constitutive model is full viscoelastic-plastic, captures strain rate dependency, temperature dependency and relaxation. **Figure 8-6** shows the capability of the model to capture large plastic deformation, necking and drawing.

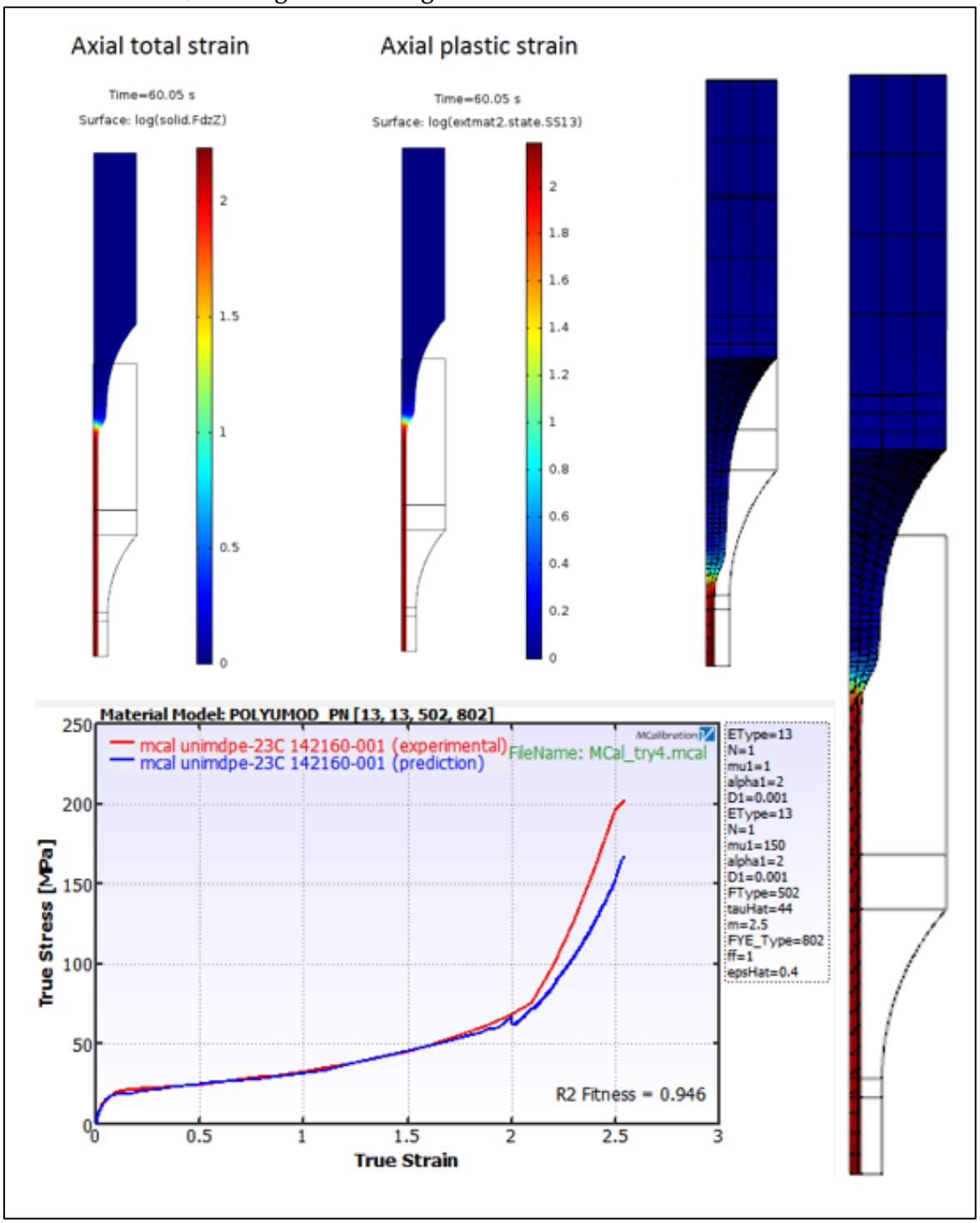


Figure 8-6. Constitutive model for FEM analysis of unimodal MDPE

Figure 8-7 shows the true stress / true strain and engineering stress / strain curves output by the constitutive model. In the engineering stress-strain curve we see strain hardening from about 800% engineering strain. In **Figure 8-8** we show how to identify the onset of strain hardening on the true stress strain curve.

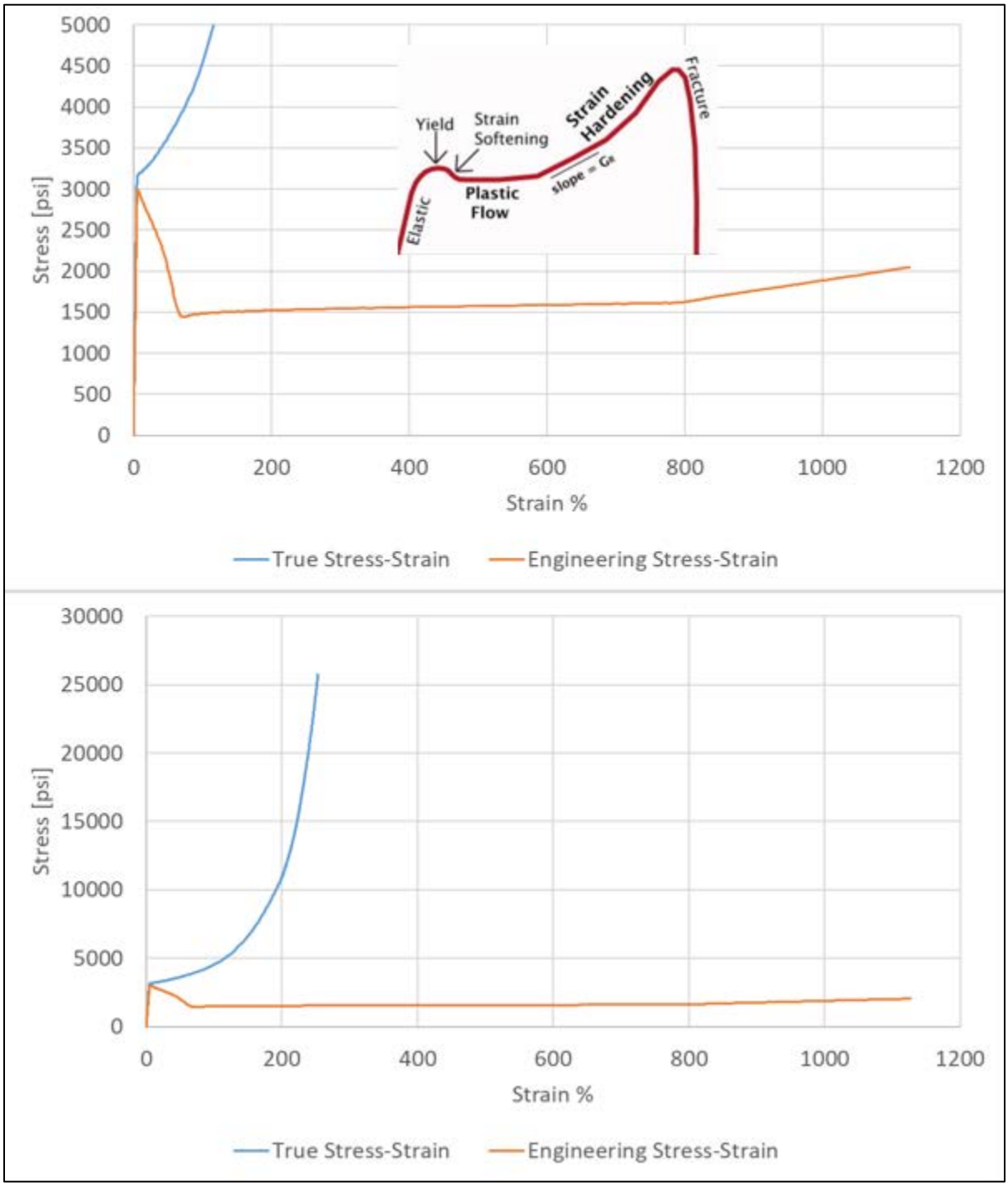


Figure 8-7. Engineering and True Stress / True Strain curves from the constitutive model

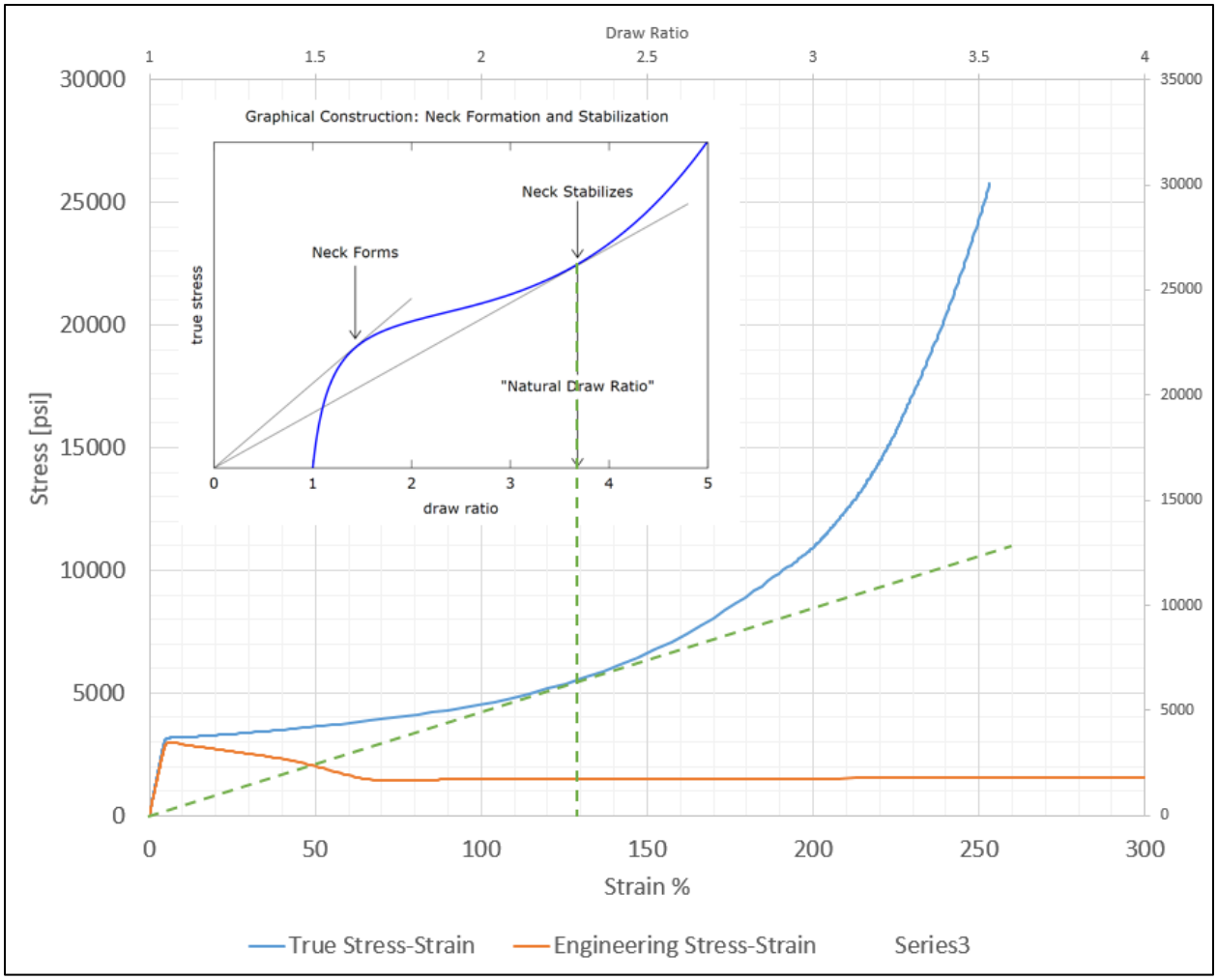


Figure 8-8. Draw ratio and onset of strain hardening

The damage propagation resistance of different materials will be strongly dependent on the strain hardening characteristics of the material and the dependence of the critical break strain as a function of stress triaxiality. There are advanced damage propagation models that address these concepts [90-94]. Detailed discussion of these approaches is beyond the scope of this project, but GTI is actively exploring ways to calibrate these models to polyethylene materials and how to incorporate them into time to failure simulations for various damage propagation regimes.

Figure 8-9 shows the first principal stress that can barely be seen at the image scale and the stress triaxiality contours. The stress triaxiality is a measure of the 2-D constraint. At higher stress triaxiality more plasticity will be experienced and damage will propagate faster.

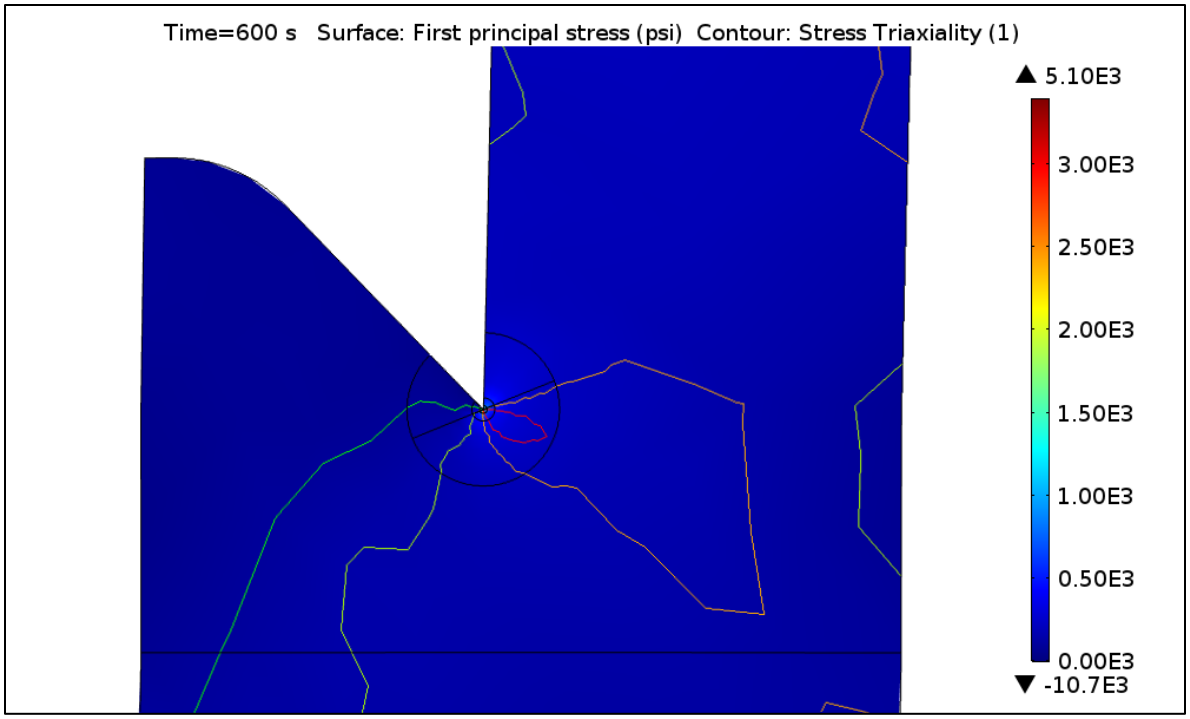


Figure 8-9. FEM analysis of the stress regime at the UTV notch tip

Stress triaxiality is defined as the ratio of hydrostatic pressure, or mean stress, to the von Mises equivalent stress.

$$\frac{\sigma_h}{\sigma_{eqv}} = TF = \frac{1/3(\sigma_1+\sigma_2+\sigma_3)}{\frac{1}{\sqrt{2}}\sqrt{(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2}}$$

where σ_1, σ_2 and σ_3 are the first, second and third principal stresses and

$$\frac{\sigma_1+\sigma_2+\sigma_3}{3} = \text{Hydrostatic stress}$$

$$\frac{1}{\sqrt{2}}\sqrt{(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2} = \text{Equivalent von Mises stress}$$

https://www.quora.com/What-is-stress-triaxiality accessed 6/25/2017

The principal direction of the stress triaxiality indicates the direction in which the crack or damage will propagate. **Figure 8-10** gives a better feel for the degree of constraint at the notch tip. The actual stress field exists in a small volume, while the stress triaxiality field exists at larger scale and dictates the direction of propagation.

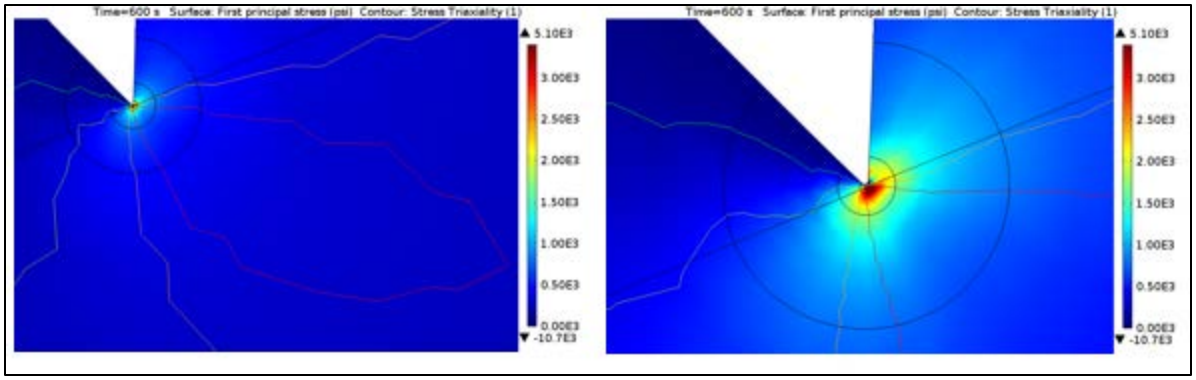


Figure 8-10. Zooming in to better visualize the notch tip stress field

Intact UTV specimens that failed by suspected SCG from the notch tip were examined by high resolution 3D X-Ray Computed Tomography capable of resolving volumes around one cubic micron. **Figure 8-11** shows an image of a cross-section through the UTV that closely matches the FEM output shown in .

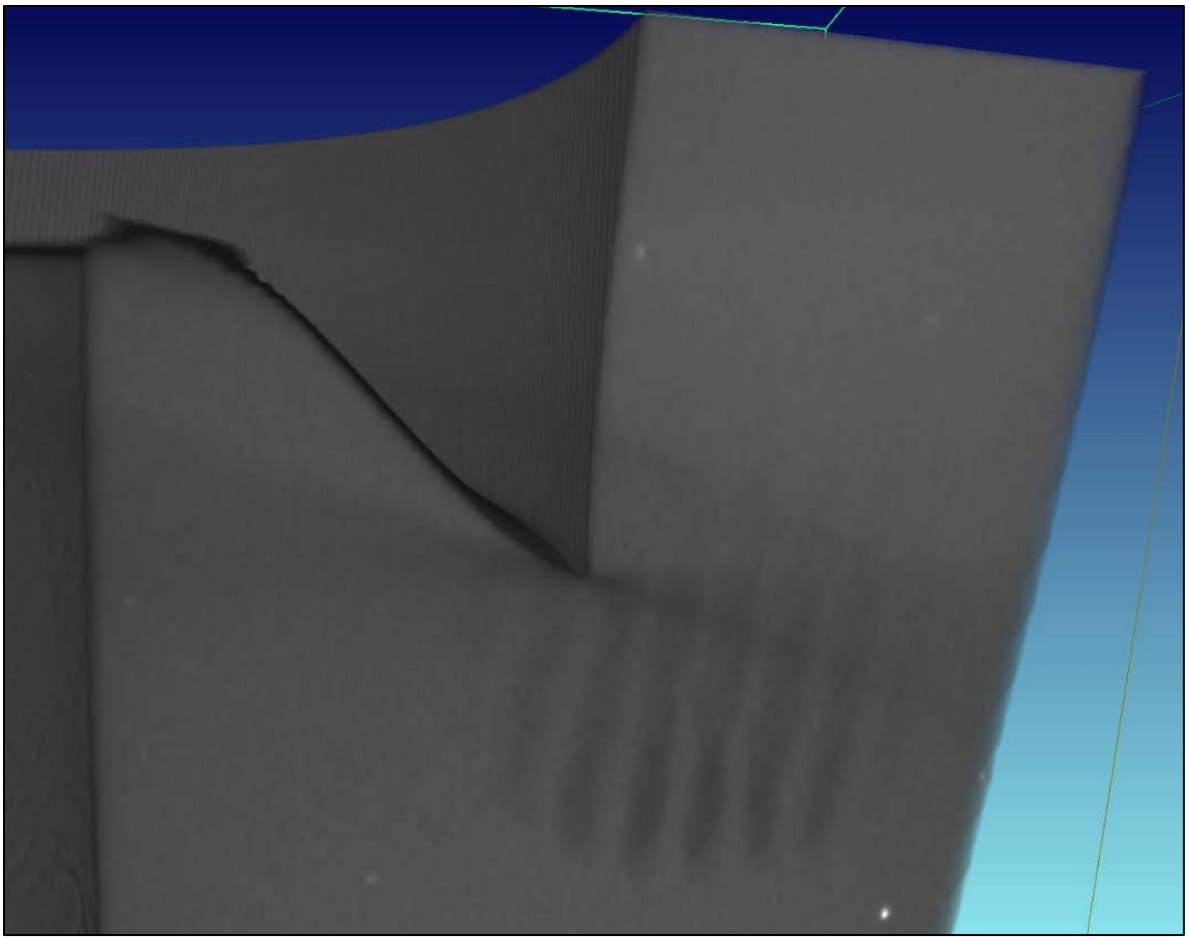


Figure 8-11. Computerized X-Ray Tomography image of damage propagation from the notch tip in a first generation unimodal MDPE UTV

Figure 8-12 shows a view, orthogonal to the previous view, where we can clearly the consistency of the molded stress riser in the UTV. There is consistent and symmetric damage propagation into the specimen wall from the notch tip.

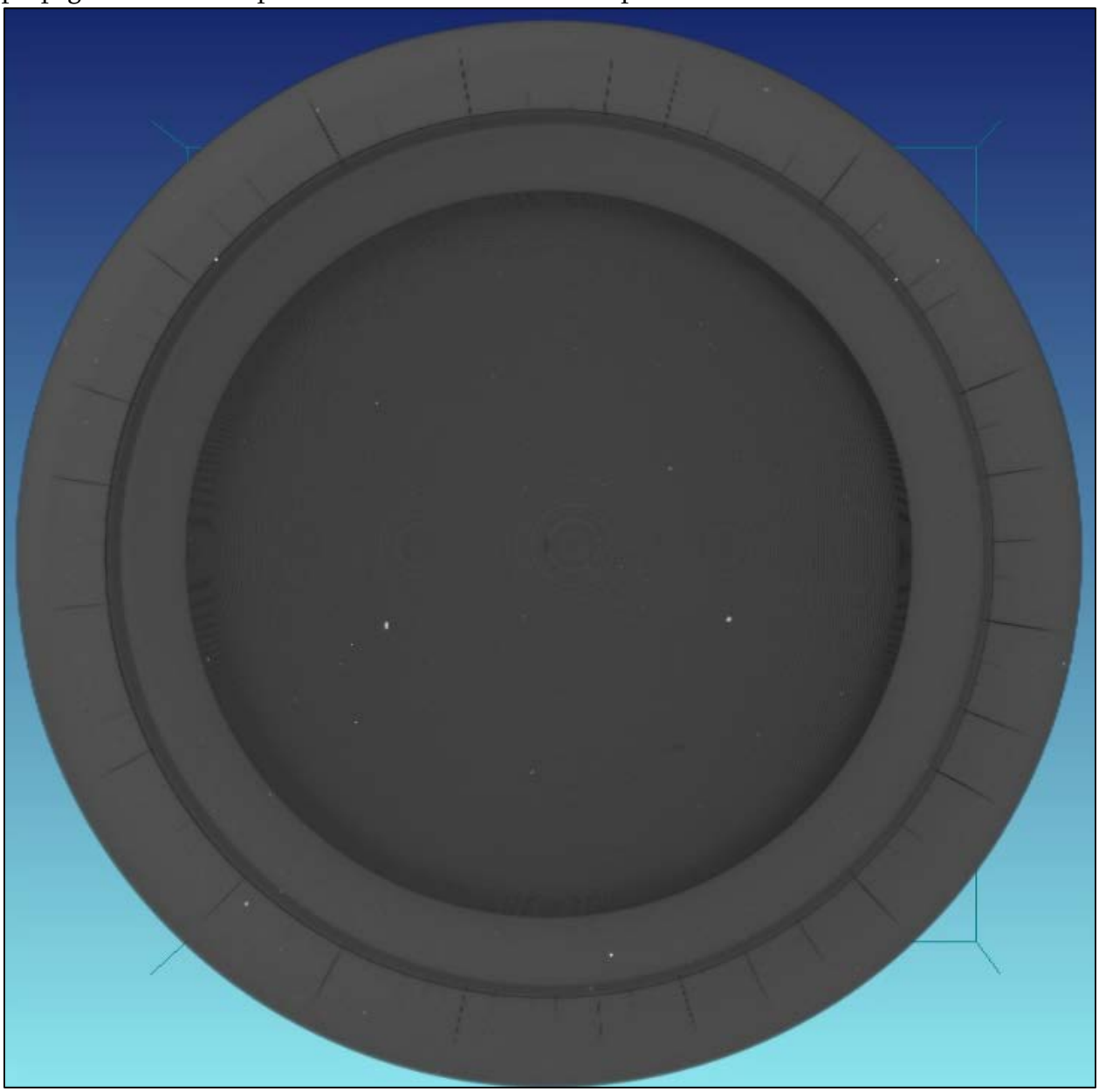


Figure 8-12. Computerized X-Ray Tomography section through UTV showing multiple damage fronts propagating radially outwards from the notch tip

We have now seen that we can clearly predict the locus and direction of damage propagation in polyethylene bodies. We will now look at fracture surfaces from damage propagation in a UTV with no stress riser, i.e. the damage propagated from an intrinsic defect in the wall of the specimen **Figure 8-13** to **Figure 8-19**.

Figure 8-13. Specimen outer wall showing breach

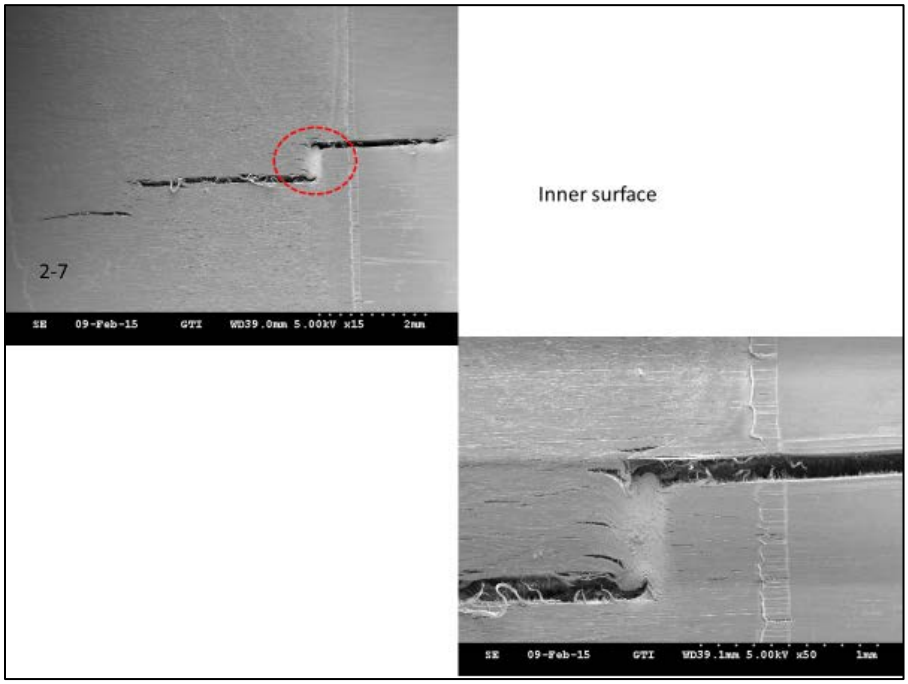


Figure 8-14. Inner surface showing stress cracking

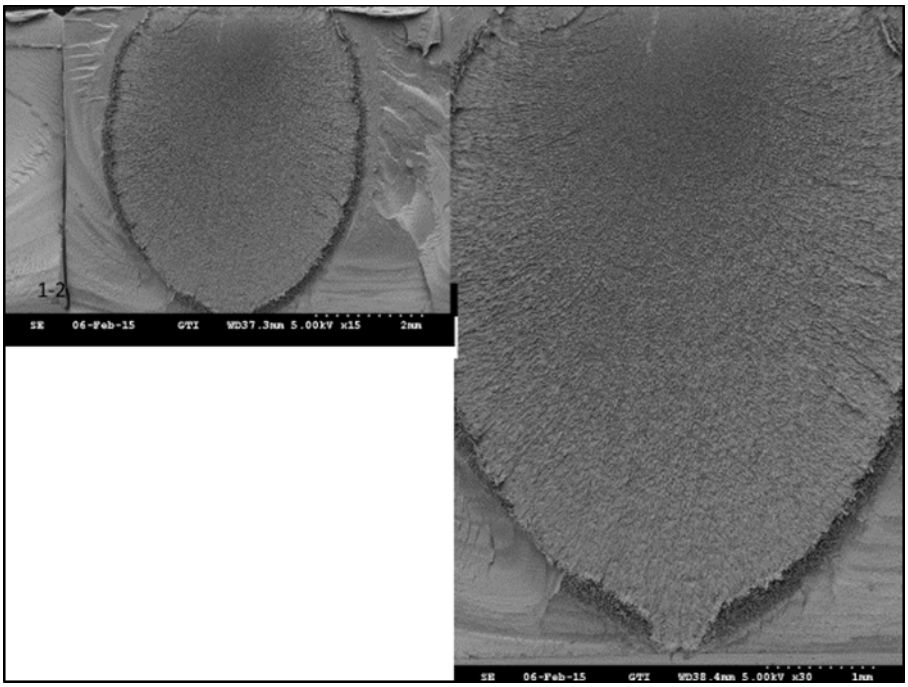


Figure 8-15. Fracture surface showing inner wall at top and outer wall at bottom

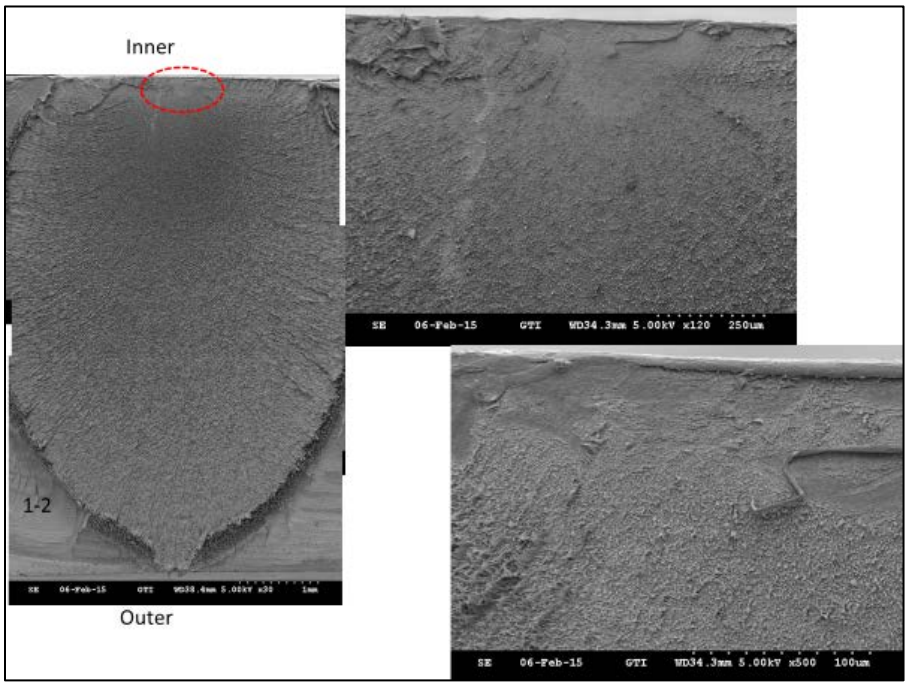


Figure 8-16. Detail of fracture surface – inner region

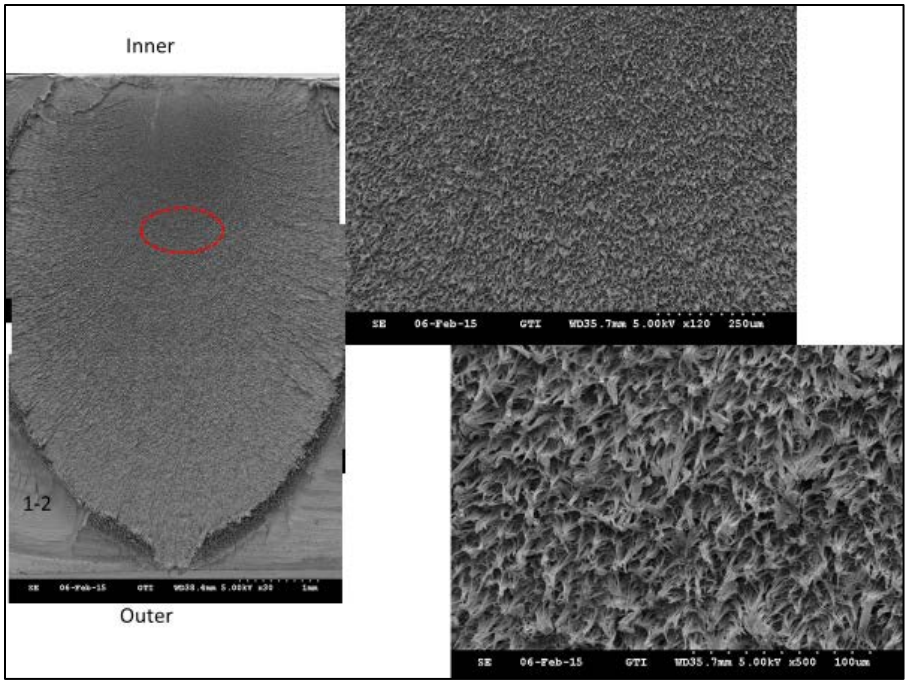


Figure 8-17. Detail of fracture surface – inner third region

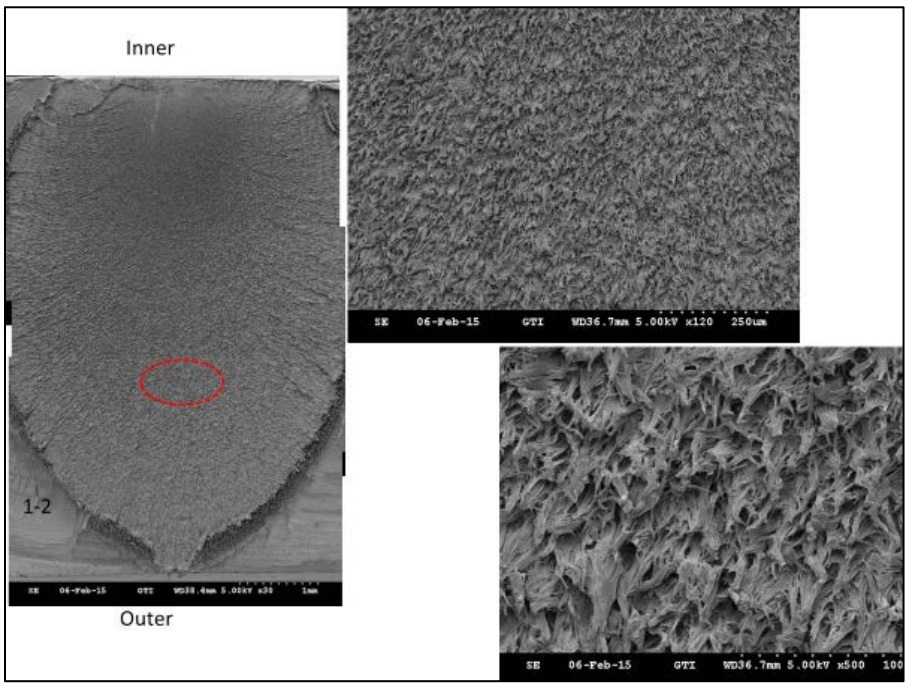


Figure 8-18. Detail of fracture surface – outer third region

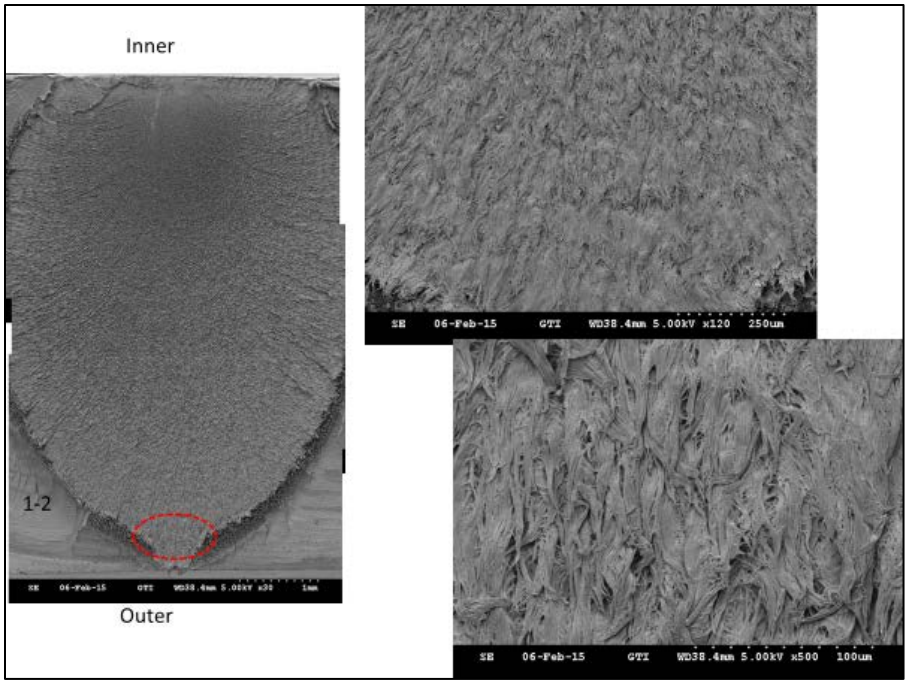


Figure 8-19. Detail of fracture surface – wall breach region

In reviewing **Figure 8-13** to **Figure 8-19** we note first that on the inner surface we see cracks with a large opening on the surface and cracks with a narrower opening on the surface. The size of this opening is a function of how far into the wall the crack has progressed. From simple geometry, we can understand that the further the crack front progresses into the wall, the less constrained it is geometrically. When we look at the fracture morphology we see a transition from a relatively smooth fracture surface with short, finer fibrils to increasingly larger and longer fibrils. As we move into the wall-breach region we see relatively large volumes that have been "torn" with "smeared" looking "leaves" that have clearly undergone large ductile deformation.

There are two factors at play in this progression:

1. The SIF increases with increasing crack depth
2. Geometric constraint is lost as the crack depth increases
3.

These two phenomena are graphically illustrated in **Figure 8-20** and **Figure 8-21** that show results of an FEM analysis using a simple constitutive model on an externally notched pipe conforming to the CPF test specimen geometry. We see that as the depth of damage penetration into the pipe wall increases the tip stress increases dramatically and triaxiality decreases sharply until it reaches a lower bound a little less than 0.4. A state of uniaxial tension has a stress triaxiality of 1/3. We can use **Figure 8-21** to determine, as a function of

internal pressure, at what depth of damage propagation we achieve ligament stress (triaxiality approaching 1/3) at the damage tip.

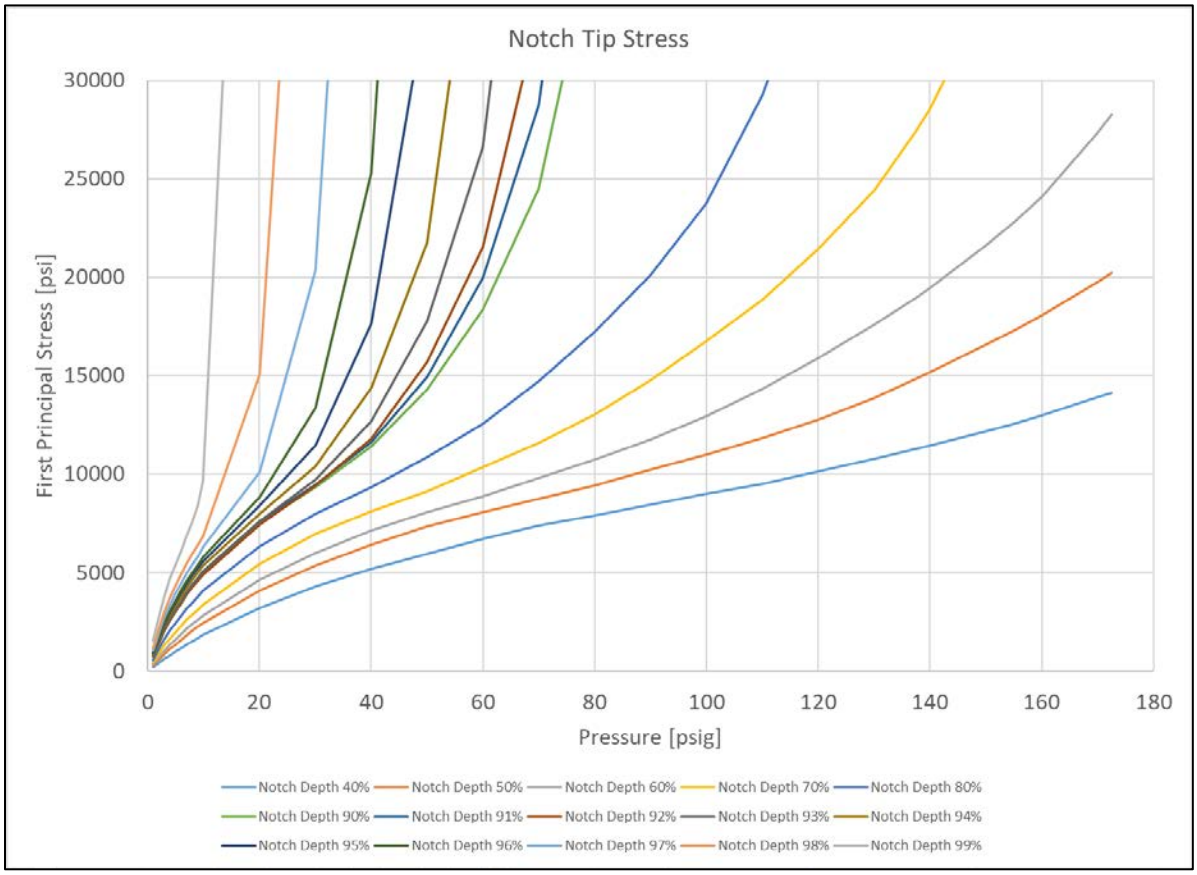


Figure 8-20. Increasing damage tip stress as a function of test pressure and depth of damage in an externally notched pipe

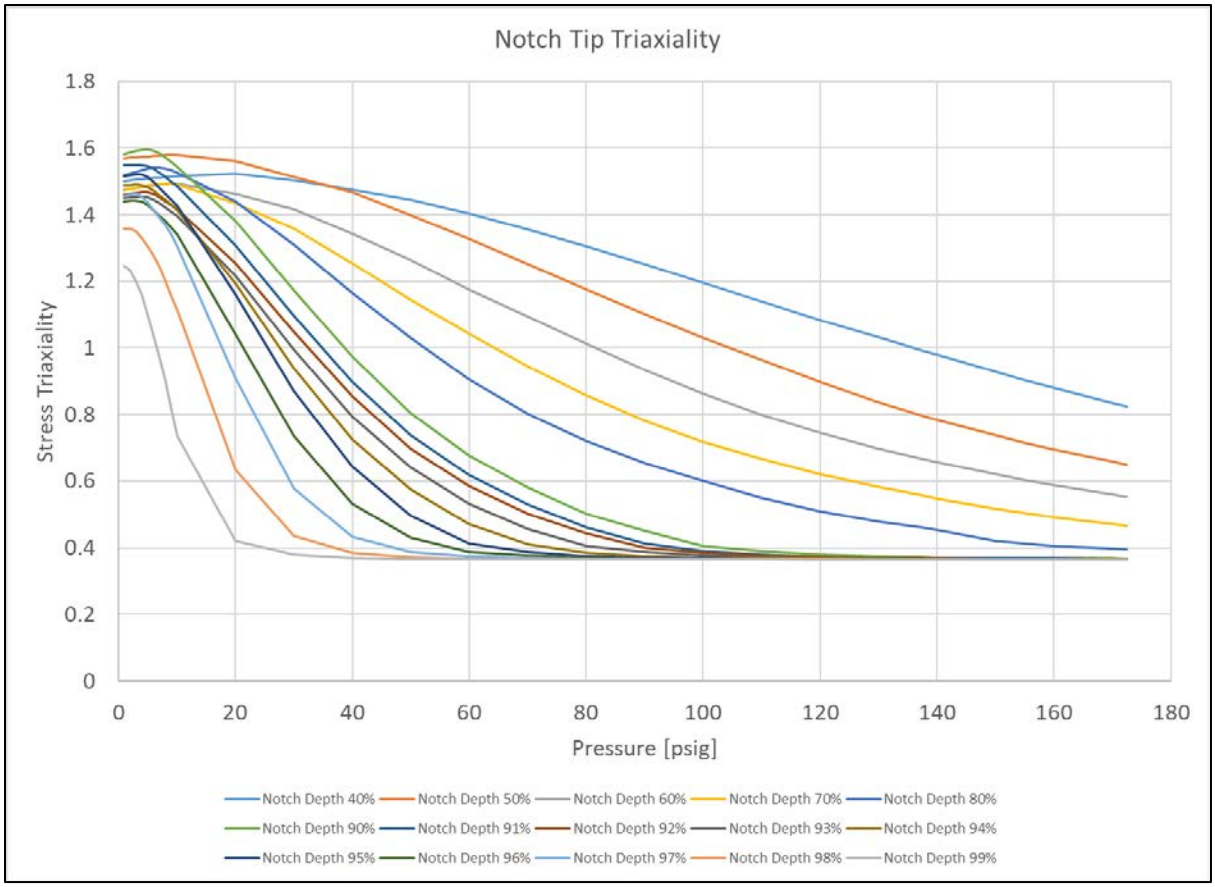


Figure 8-21. Decreasing triaxiality (loss of constraint) at the damage tip as a function of test pressure and depth of damage in an externally notched pipe

We will now address, in some detail, void formation ahead of the damage zone that occurs at high stress triaxiality (these voids give rise to fibrils). Voids form at stress triaxiality levels greater than 1 as shown in **Figure 8-22** and **Figure 8-23**.

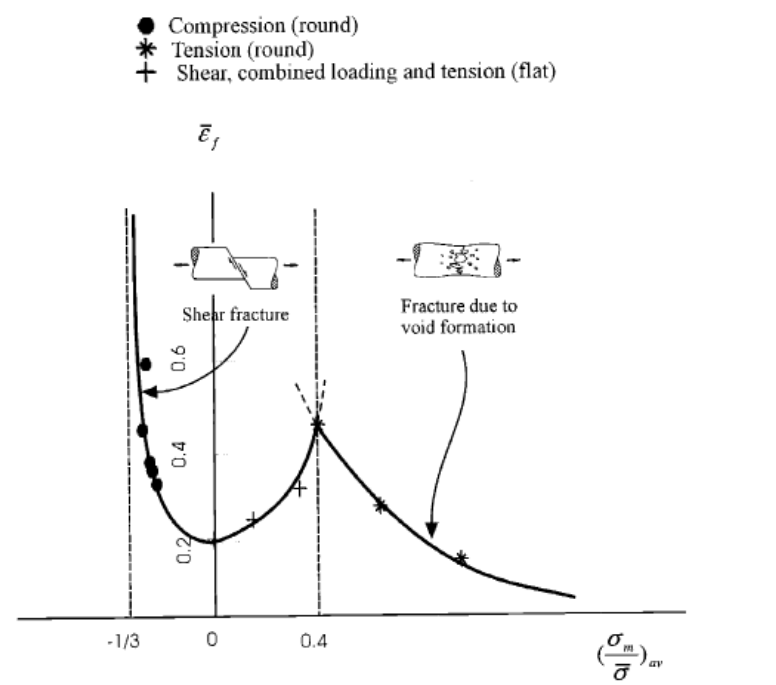


Fig. 5.20 Dependence of the equivalent strain to crack formation on the stress triaxiality

Figure 8-22. Dependence of critical strain at break and fracture mode of stress triaxiality per Bao [90] (PhD Thesis, MIT 2003)

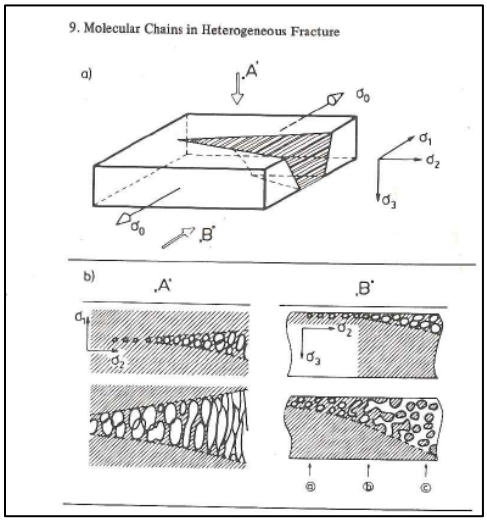


Figure 8-23. Void and fibril formation in polymers under stress triaxiality. Kausch, Polymer Fracture, Fig 9.17 a, p 340 [95]

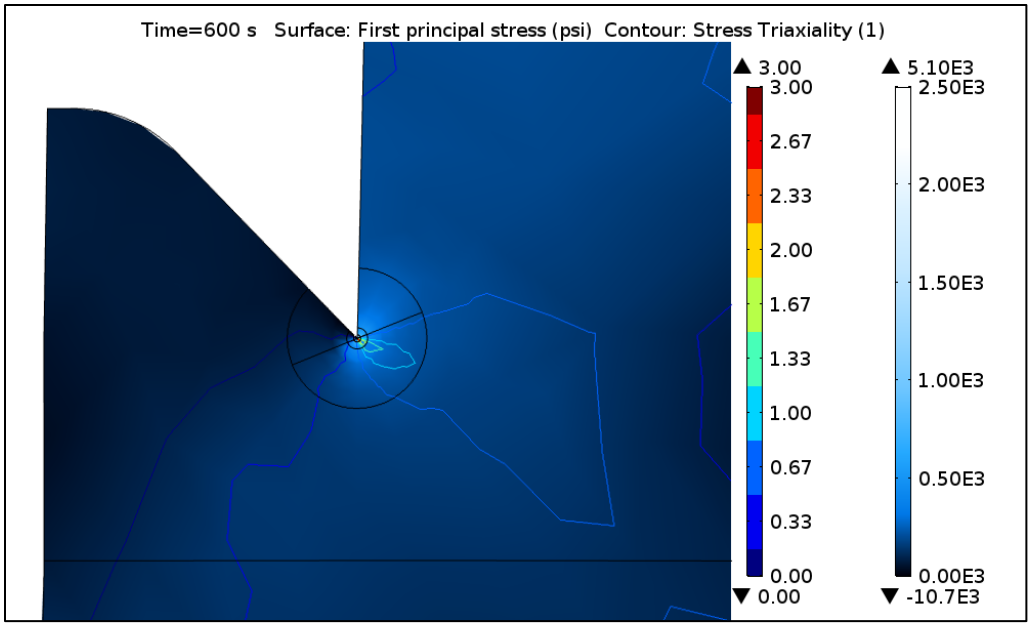


Figure 8-24. Stress triaxiality contours around UTV notch. Outer closed contour is for triaxiality = 1/3, and demarks the region where hoop stress dominates (outside the closed contour).

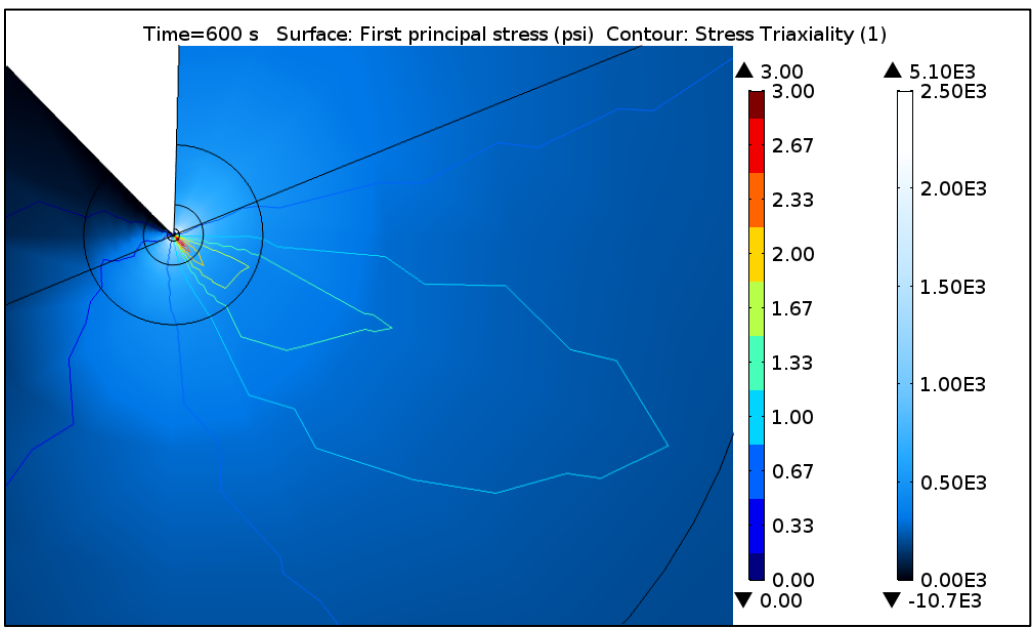


Figure 8-25. Stress triaxiality contours around UTV notch. Outer closed contour is for triaxiality = 1, and demarks the region where void formation is possible (inside the closed contour).

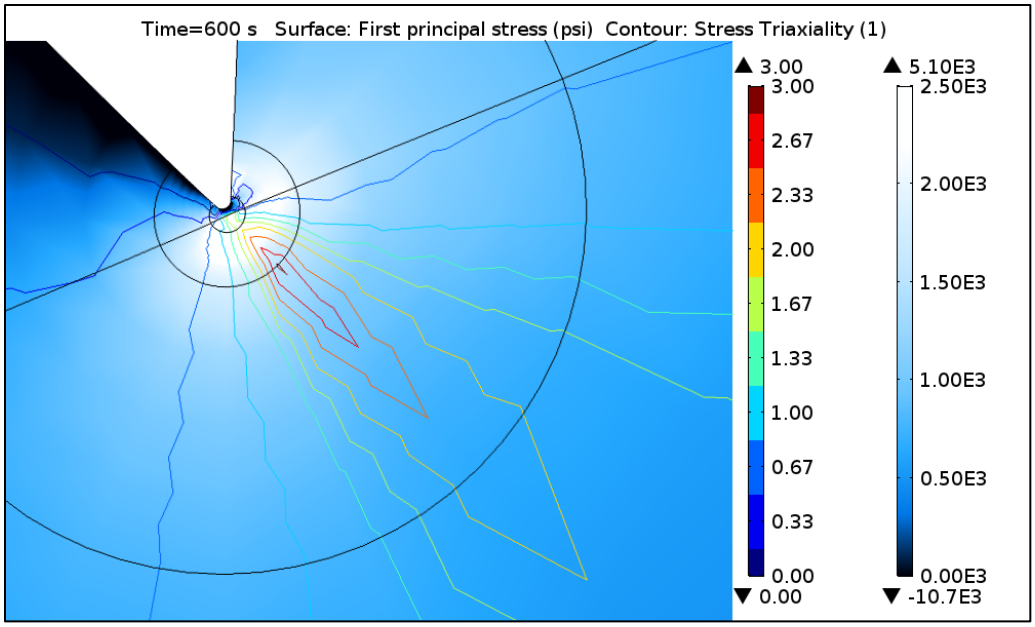


Figure 8-26. Stress triaxiality contours around UTV notch. innermost closed contour is for triaxiality = 3, and demarks the region where void formation is highly likely to occur (inside the closed contour).

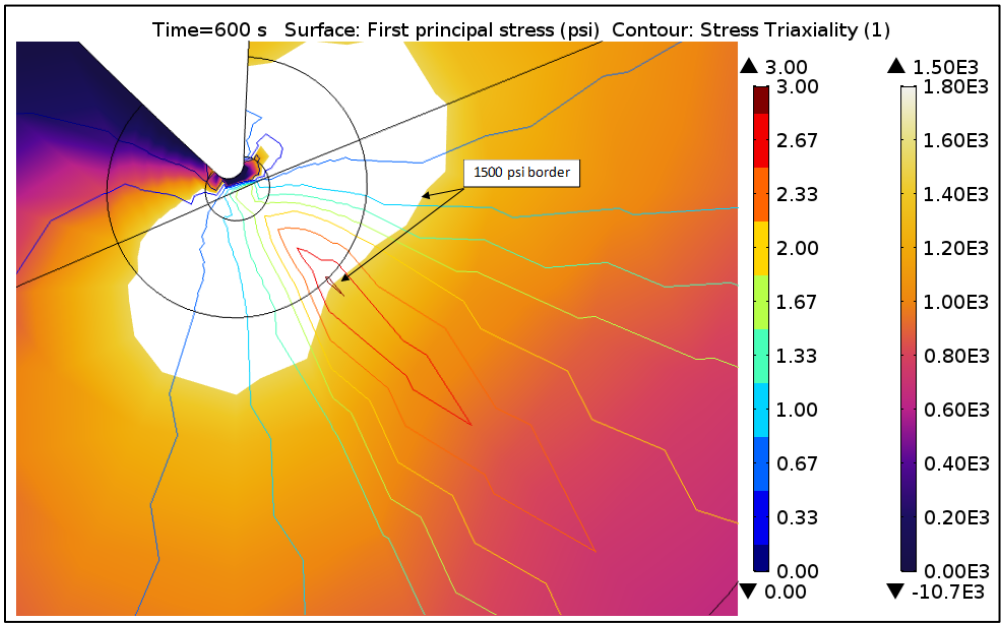


Figure 8-27. Same as Figure 8-26 – accentuating the stress levels at the maximum triaxiality region

Running a viscoelastic/plastic FEM analysis on the externally notched pipe specimens used in overpressure testing conducted by GTI we see in **Figure 8-28** to **Figure 8-30** that we have potential for SCG at a damage penetration depth of 40% and 550 psi hoop stress. At a damage penetration depth of 70% there is already little likelihood of SCG.

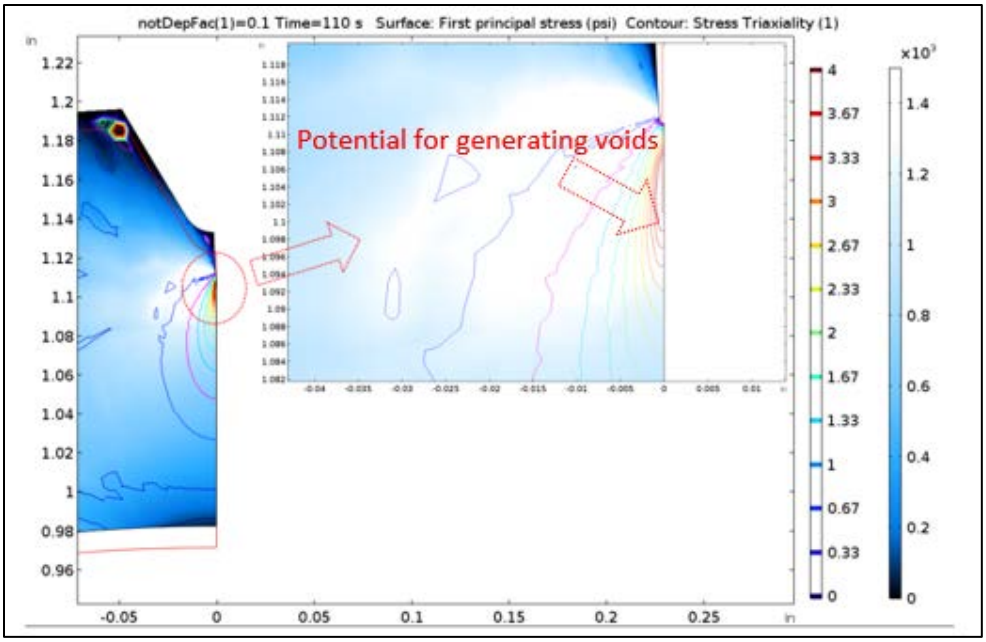


Figure 8-28. External Notch, 40% depth, 110 psig internal pressure (550 psi hoop stress)

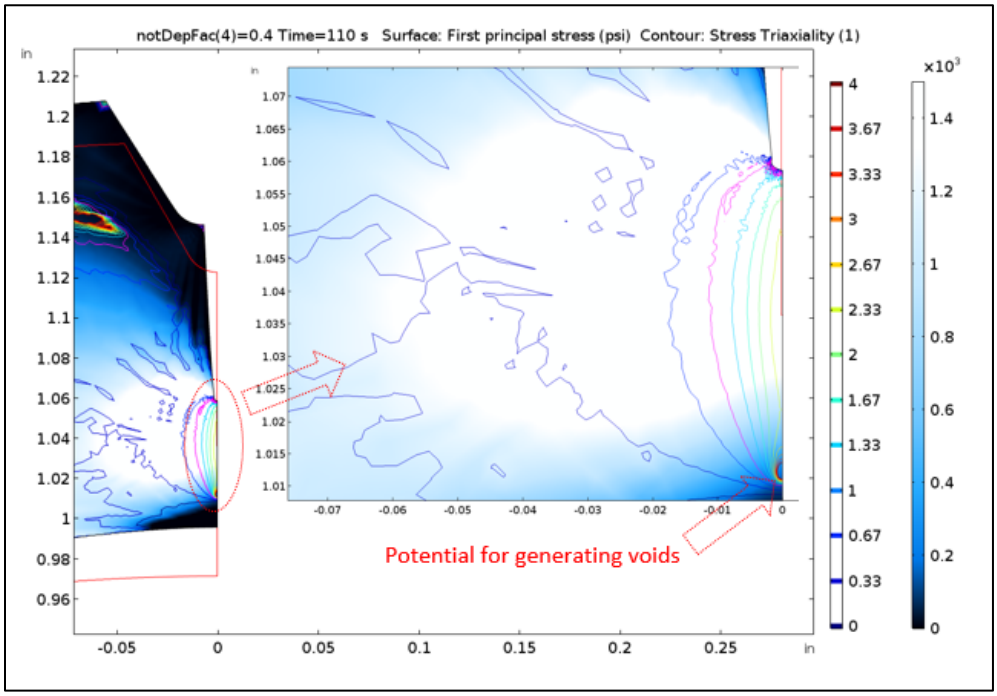


Figure 8-29. External Notch, 70% depth, 110 psig internal pressure (550 psi hoop stress)

All the FEM analyses were run for polyethylene at 23°C (73°F).

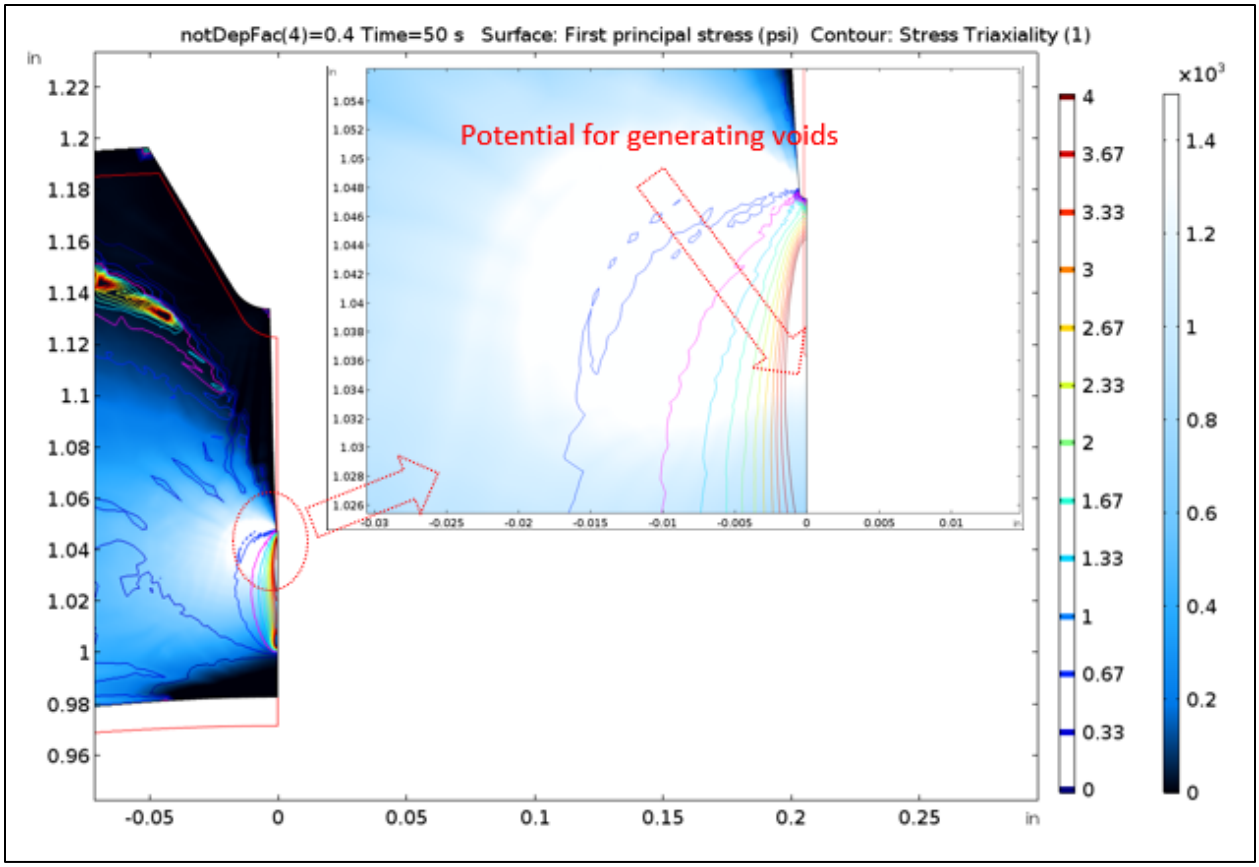


Figure 8-30. External Notch, 70% depth, 50 psig internal pressure (250 psi hoop stress)

Now that we have a rudimentary understanding of why voids and fibrils form ahead of the notch tip, we need to recognize that the behavior of the individual fibrils will exhibit some stochastic variation due to random variations in material properties in each fibril.

We will first look at variation in ultimate strain at break under displacement controlled yielding.

In **Figure 8-31** we see actual variation in strain at break in similar specimens cut from pipe. In **Figure 8-32** we show the energy density for each of the tensile specimens. These two plots make the connection between critical strain at break and the toughness of the material i.e. how much energy it absorbs before fracture. The fibrils in SCG and the remaining ligament in the final ductile rupture will follow these curves on their path to rupture.

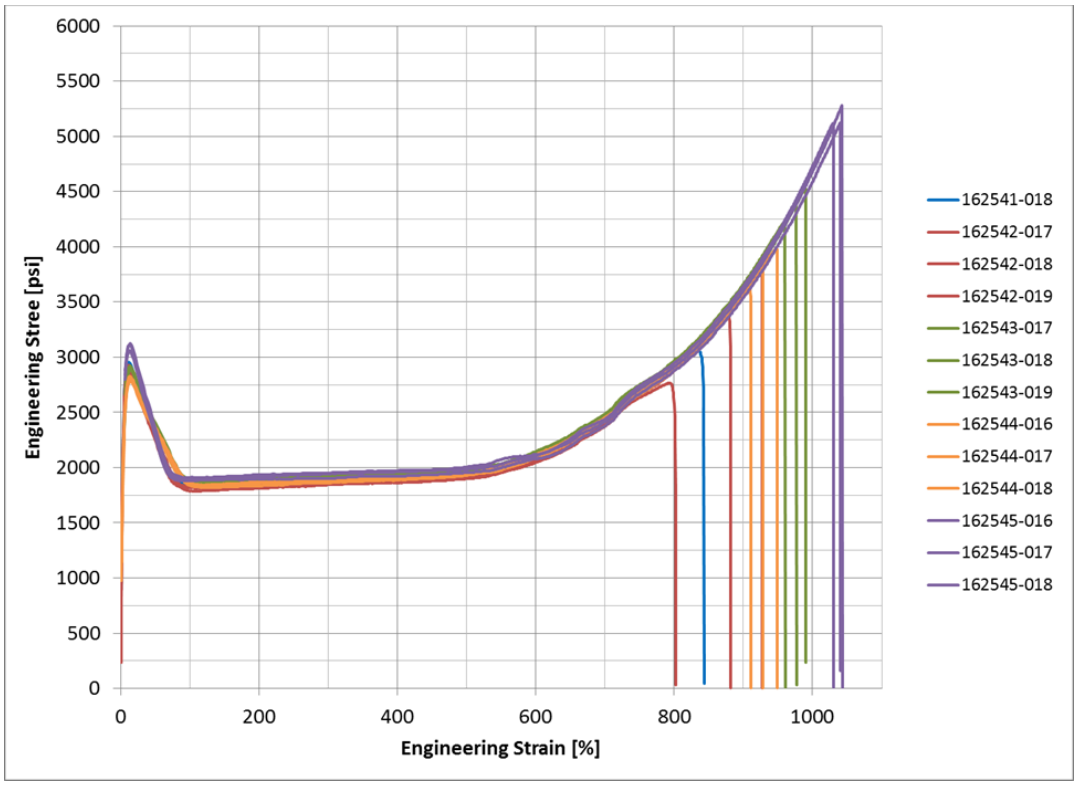


Figure 8-31. Variations in critical strain at break in multiple unimodal MDPE specimens, cut from the same pipe, under the same nominal tensile test conditions

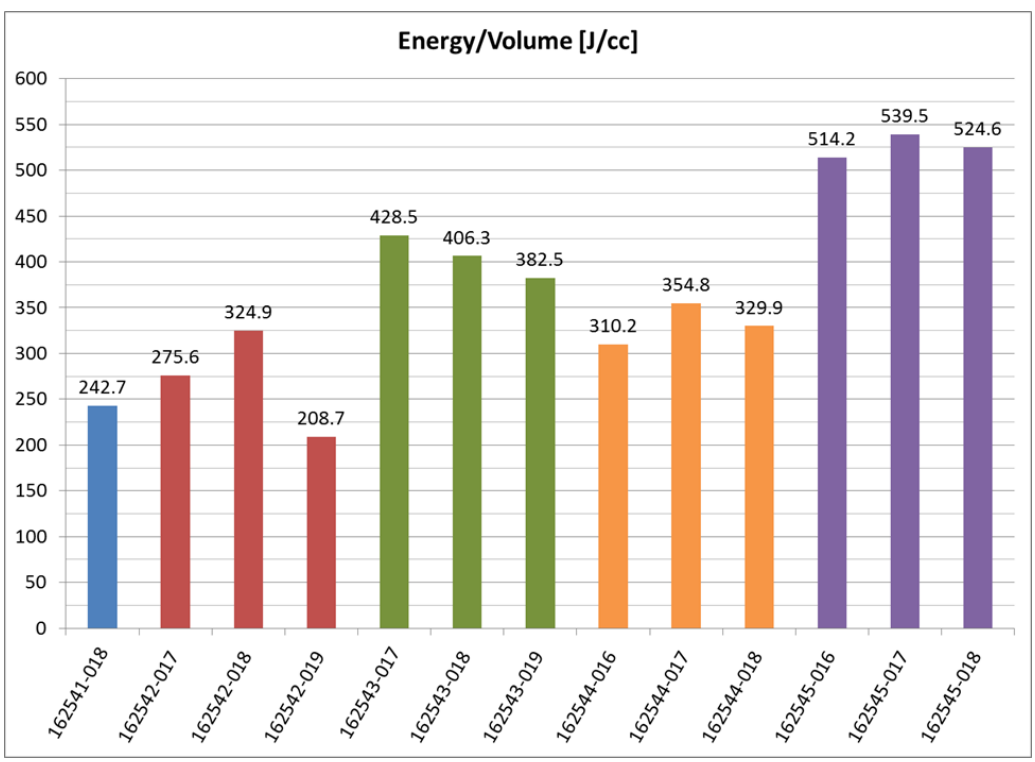


Figure 8-32. Variation in the energy density for the tensile tests shown in Figure 8-31

Figure 8-33 and **Figure 8-34** show load controlled creep results at 90°C and 5 MPa (725 psi) nominal stress. The ASTM D638 Type V dog bones, punched from the same pipe were left on test for 4313 hours (equivalent to > 400 years and > 10 MPa (1450 psi) stress @ 20°C. The remarkable result is the spread in creep rates where there is an order of magnitude difference between the highest and lowest extensions over this time. Upon displacement controlled pull to break after 4313 hours of force controlled creep, both specimens converged to the same force displacement curve and failed at peak loads and displacements within 3% of their average value.

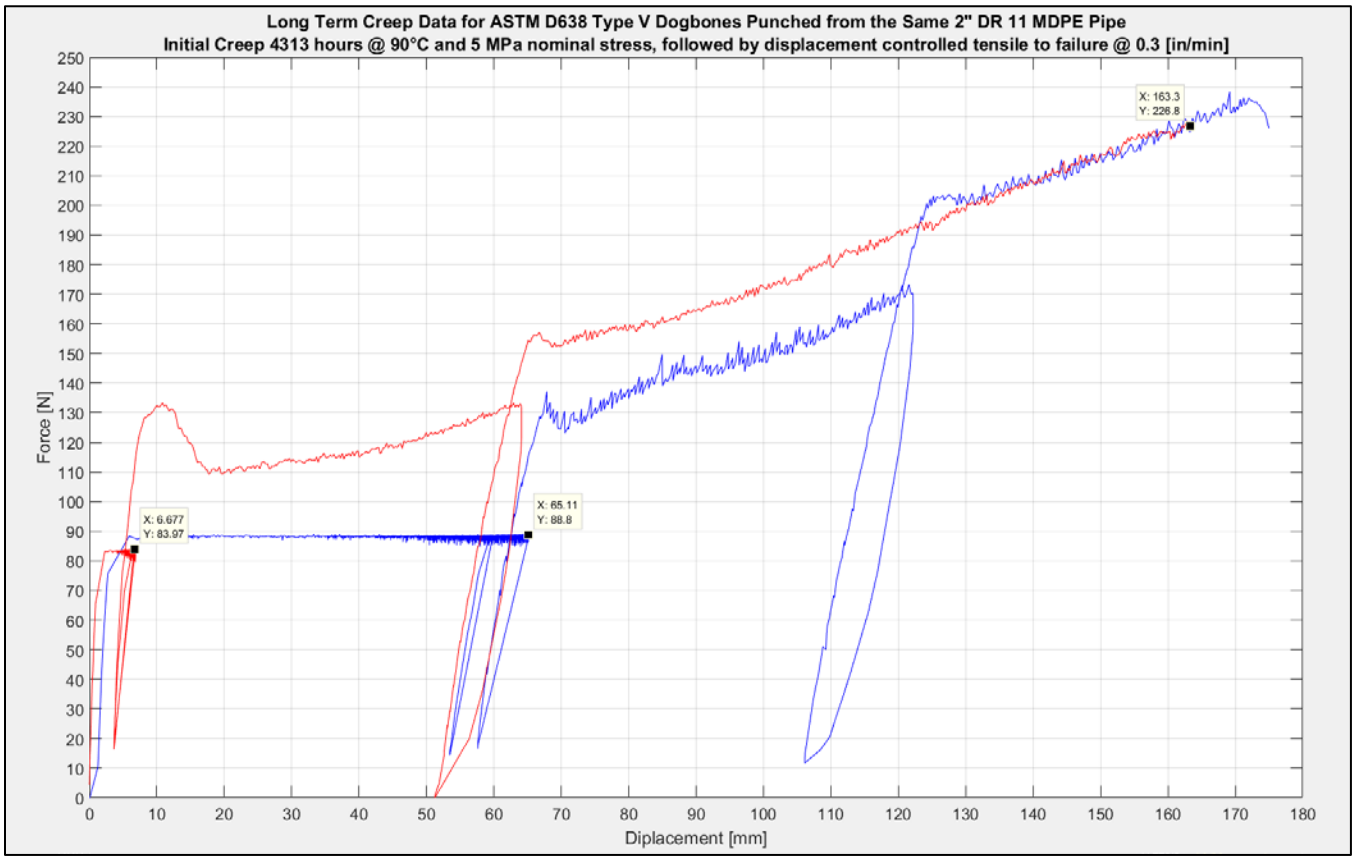


Figure 8-33. MDPE - comparison of creep rates under load controlled long-term tensile creep @ 90°C, 5 MPa

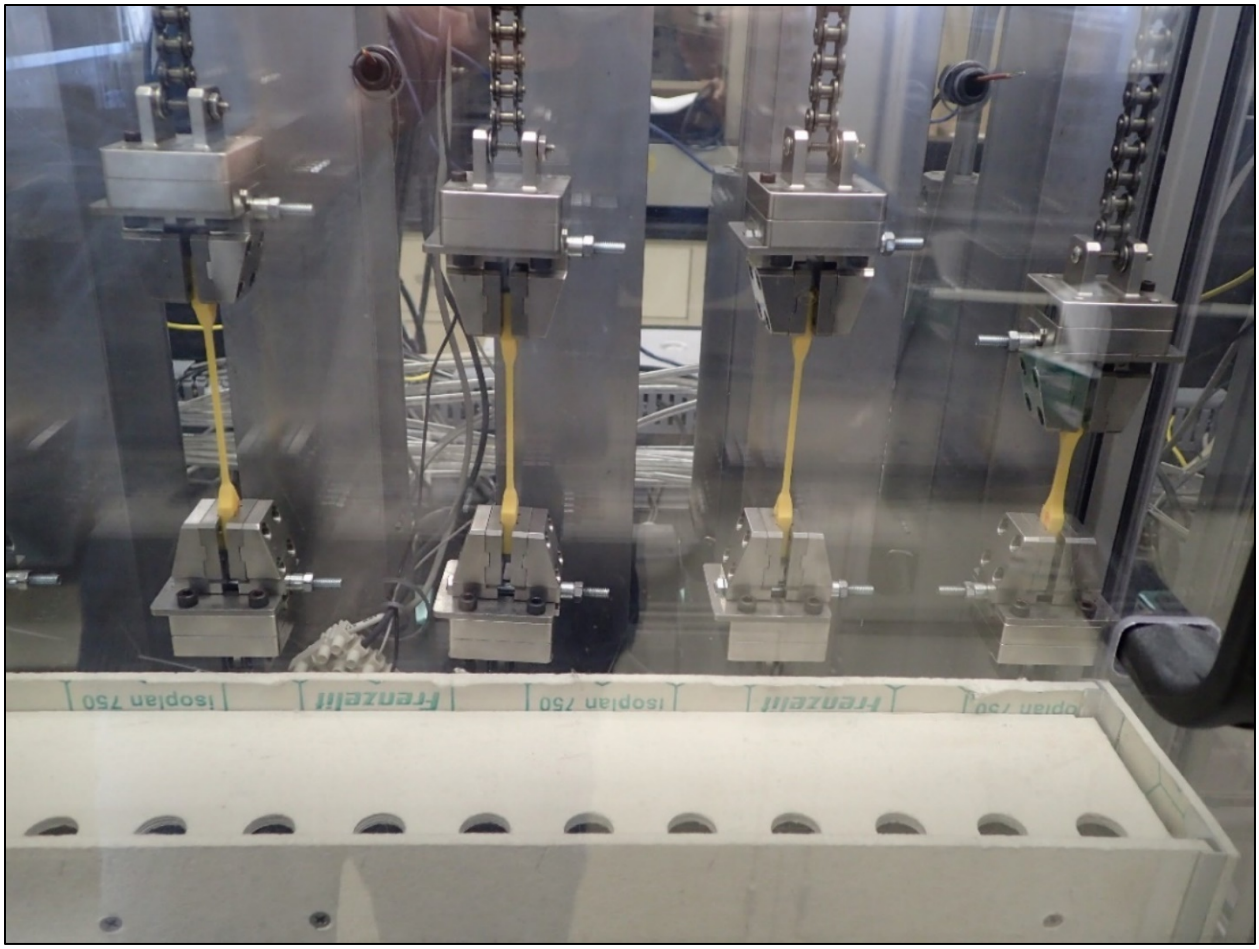


Figure 8-34. Left specimen blue curve above, Right specimen red curve above

The creep data from 27 ASTM D638 Type V dog bone specimens punched from 3 similar pipes, exhumed from service in a gas distribution system is presented in **Table 8-1**. The data is color coded per pipe. The sixth digit of the LIMS number identifies the pipe and the last three digits the specimen number punched from the pipe.

Table 8-1. Creep Data for 2" SDR 11 MDPE Gas Distribution Pipe. Color Coded Per Pipe (3 pipes with multiple specimens from each)

Station	LIMS	State	TempdegC	AvgTdegC	AvgTK	AvgStressMPa	ReferenceStrain	EnergyVolJcc
'1-1'	'162541-025'	'Creep @ 4313 h'	90	89.8	363.0	5.0	90.4	5653.6
'1-2'	'162541-026'	'Creep @ 4313 h'	90	89.8	363.0	5.0	75.4	5293.3
'1-3'	'162541-027'	'Creep @ 4313 h'	90	89.8	363.0	5.0	756.3	39202.8
'1-4'	'162542-026'	'Creep @ 4313 h'	90	89.8	363.0	4.7	721.5	35648.9
'1-5'	'162542-027'	'Creep @ 4313 h'	90	89.8	363.0	4.7	345.5	18145.4
'1-6'	'162542-028'	'Creep @ 4313 h'	90	89.8	363.0	4.7	569.0	28329.2
'1-7'	'162543-026'	'Creep @ 4313 h'	90	89.8	363.0	4.8	66.6	3695.5
'1-8'	'162543-027'	'Creep @ 4313 h'	90	89.8	363.0	4.8	64.4	3829.2

Station	LIMS	State	TempdegC	AvgTdegC	AvgTK	AvgStressMPa	ReferenceStrain	EnergyVolJcc
'1-9'	'162543-028'	'Creep @ 4313 h'	90	89.8	363.0	4.8	49.2	3086.4
'1-1'	'162541-025'	'Break'	90	89.8	363.0		1049.7	69666.0
'1-2'	'162541-026'	'Break'	90	89.8	363.0		2426.0	241034.1
'1-3'	'162541-027'	'Break'	90	89.8	363.0		2300.8	197375.2
'1-4'	'162542-026'	'Break'	90	89.8	363.0		2444.2	212809.8
'1-5'	'162542-027'	'Break'	90	89.8	363.0		2580.6	234933.0
'1-6'	'162542-028'	'Break'	90	89.8	363.0		1967.2	158845.9
'1-7'	'162543-026'	'Break'	90	89.8	363.0		2156.1	210488.4
'1-8'	'162543-027'	'Break'	90	89.8	363.0		2219.9	217476.7
'1-9'	'162543-028'	'Break'	90	89.8	363.0		2150.9	202415.3
'2-1'	'162541-022'	'Creep @ 4313 h'	60	66.1	339.2	6.7	654.3	47106.5
'2-2'	'162541-023'	'Creep @ 4313 h'	60	66.1	339.2	6.7	507.9	38271.9
'4-1'	'162541-024'	'Creep @ 4313 h'	60	60.7	333.9	6.7	656.3	46530.7
'2-4'	'162542-023'	'Creep @ 4313 h'	60	66.1	339.2	6.3	606.2	41137.5
'2-5'	'162542-024'	'Creep @ 4313 h'	60	66.1	339.2	6.3	249.5	18745.6
'2-6'	'162542-025'	'Creep @ 4313 h'	60	66.1	339.2	6.3	647.1	43584.3
'2-7'	'162543-023'	'Creep @ 4313 h'	60	66.1	339.2	6.4	514.0	36334.1
'2-8'	'162543-024'	'Creep @ 4313 h'	60	66.1	339.2	6.4	666.7	43371.4
'2-9'	'162543-025'	'Creep @ 4313 h'	60	66.1	339.2	6.4	82.3	7995.2
'2-1'	'162541-022'	'Break'	60	66.1	339.2		2216.2	255104.6
'2-2'	'162541-023'	'Break'	60	66.1	339.2		2254.4	269912.4
'4-1'	'162541-024'	'Break'	60	60.7	333.9		2112.6	236216.8
'2-4'	'162542-023'	'Break'	60	66.1	339.2		2436.1	280357.0
'2-5'	'162542-024'	'Break'	60	66.1	339.2		2431.0	285827.6
'2-6'	'162542-025'	'Break'	60	66.1	339.2		2649.6	318821.3
'2-7'	'162543-023'	'Break'	60	66.1	339.2		2206.9	245982.6
'2-8'	'162543-024'	'Break'	60	66.1	339.2		2256.8	252660.1
'2-9'	'162543-025'	'Break'	60	66.1	339.2		1933.9	209075.6
'3-1'	'162541-019'	'Creep @ 4313 h'	40	40.3	313.5	8.9	891.4	79999.9
'3-2'	'162541-020'	'Creep @ 4313 h'	40	40.3	313.5	8.9	794.3	75752.7
'3-3'	'162541-021'	'Creep @ 4313 h'	40	40.3	313.5	8.9	771.6	73467.3
'3-4'	'162542-020'	'Creep @ 4313 h'	40	40.3	313.5	8.3	684.4	62038.7
'3-5'	'162542-021'	'Creep @ 4313 h'	40	40.3	313.5	8.3	659.4	59518.0
'3-6'	'162542-022'	'Creep @ 4313 h'	40	40.3	313.5	8.3	670.0	61012.3
'3-7'	'162543-020'	'Creep @ 4313 h'	40	40.3	313.5	8.5	67.0	10128.0
'3-8'	'162543-021'	'Creep @ 4313 h'	40	40.3	313.5	8.5	594.6	55776.2
'3-9'	'162543-022'	'Creep @ 4313 h'	40	40.3	313.5	8.5	53.6	7909.8
'3-1'	'162541-019'	'Break'	40	40.3	313.5		1688.0	218607.7
'3-2'	'162541-020'	'Break'	40	40.3	313.5		2201.1	325433.9
'3-3'	'162541-021'	'Break'	40	40.3	313.5		2234.2	333458.5

Station	LIMS	State	TempdegC	AvgTdegC	AvgTK	AvgStressMPa	ReferenceStrain	EnergyVolJcc
'3-4'	'162542-020'	'Break'	40	40.3	313.5		2367.2	360672.5
'3-5'	'162542-021'	'Break'	40	40.3	313.5		2369.7	380312.8
'3-6'	'162542-022'	'Break'	40	40.3	313.5		2164.5	311021.4
'3-7'	'162543-020'	'Break'	40	40.3	313.5		1718.4	241020.1
'3-8'	'162543-021'	'Break'	40	40.3	313.5		1639.9	215169.2
'3-9'	'162543-022'	'Break'	40	40.3	313.5		1699.8	238104.4

In **Figure 8-35** we plot the response surface for the creep data where Temperature [K] and Strain [%] are the fixed variables and energy per unit volume [J/cc] is the response. The energy per unit volume is calculated from the integral of the force displacement trace and the specimen geometry. We can clearly see that this surface is a material characteristic governed by the material density (ultimate break strength). The internal variation under force controlled creep are likely due to distributions in the rheological properties of the various molecular weight components of the resin formulation.

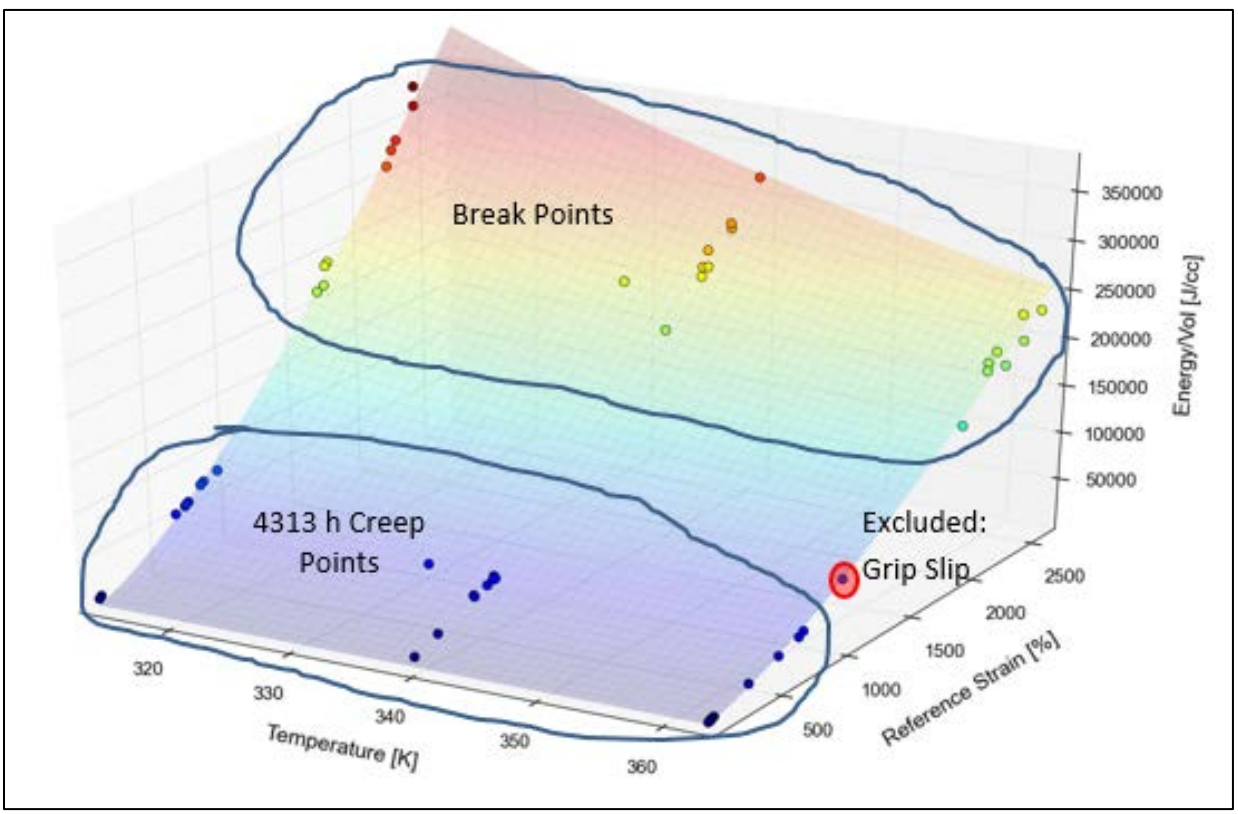


Figure 8-35. Creep Surface for MDPE

In **Figure 8-36** and **Figure 8-37** we present a graphical breakdown of the data in **Table 8-1**.

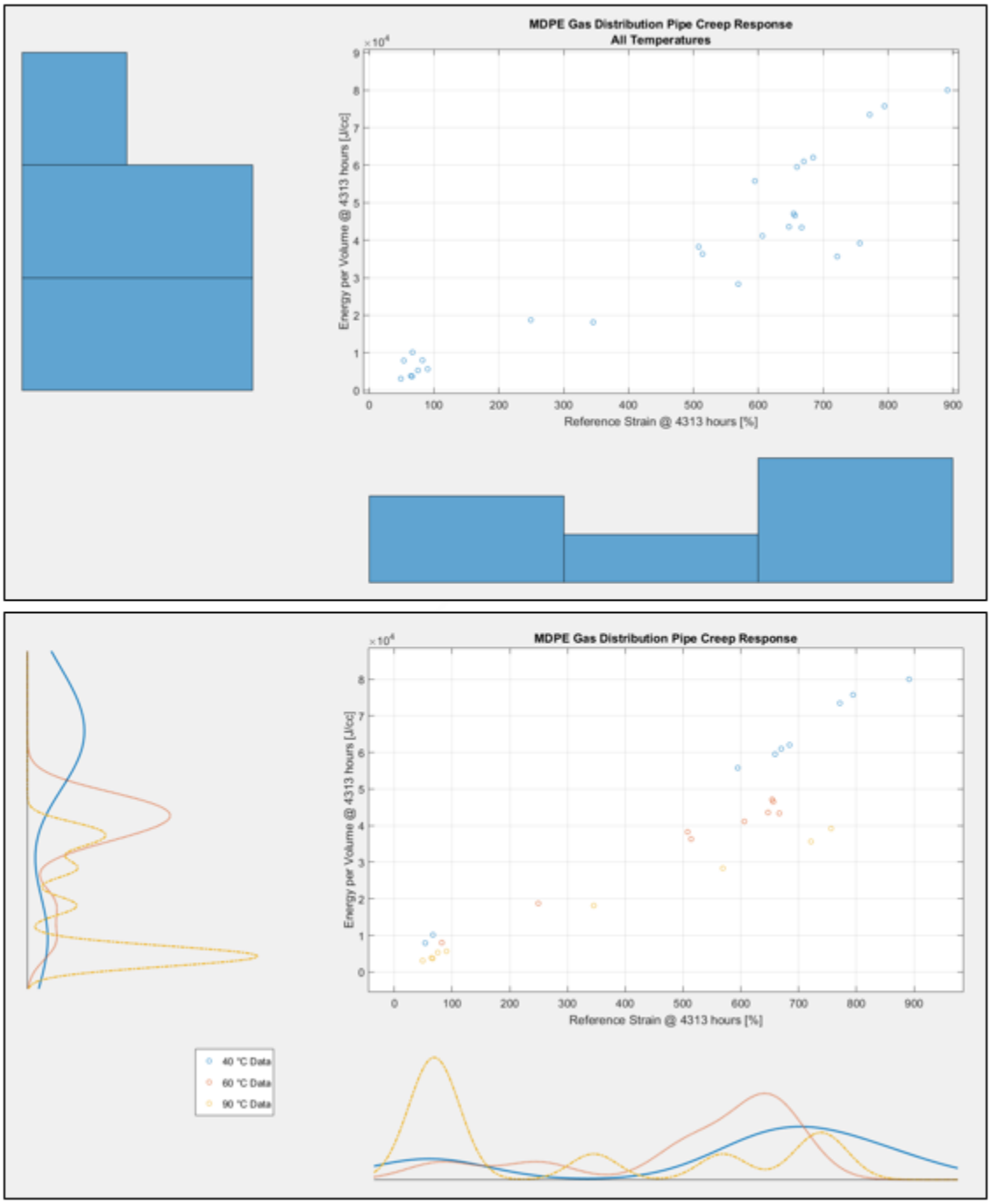


Figure 8-36. Creep Response of MDPE Pipe at 4313 h of Load Controlled Creep @ 90°C.

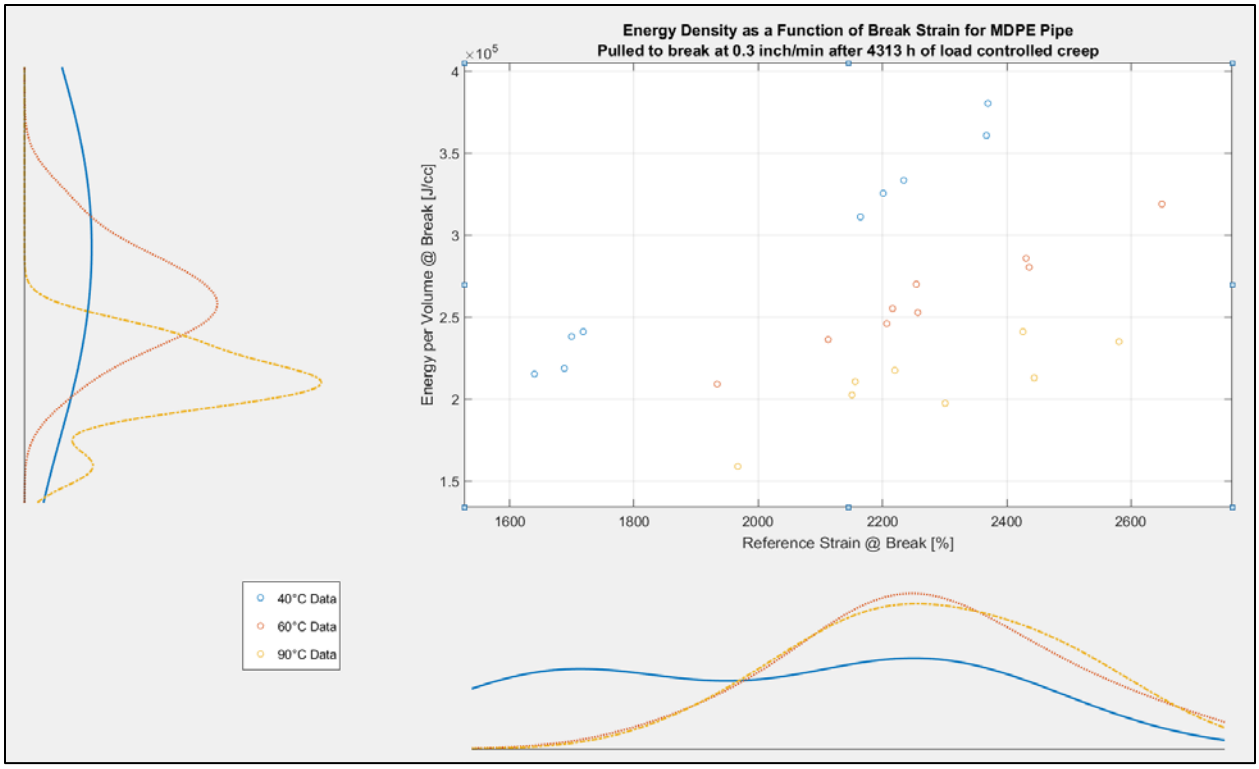


Figure 8-37. Energy Density and Strain at Break for MDPE Pipe

SCG Summary

Introducing a notch into a polyethylene pipe creates a complex situation around the notch tip that will have dependencies on geometry and material properties. Review of the stress triaxiality distribution under various geometries and internal pressure clearly indicates that SCG will only occur under relatively low nominal hoop-stress conditions. The occurrence of SCG does not indicate a new mode of failure, it is still ductile creep, but occurs in a small local volume under high geometric constraint.

Ductile Tearing Summary

GTI was able to induce ductile tearing, in polyethylene pipe, from a significant notch (30% wall thickness) during accelerated testing. Failure times of 1-10 hours were achievable at 90°C and test pressures of 150-200 psig. This would equate to test pressures of 350 – 450 psig at 68°F and failures would occur at test-times in the range 1,500-15,000 hours[12]. This is clearly impractical as field test and appears to indicate that a short-term pressure test to determine fitness for service is unlikely to succeed.

---

12 Using standard shift factors for polyethylene

8. Lever, E., *Bi-Directional Shift Factors Revisited*, in *Plastic Pipes XVII*. 2014: Chicago, IL.

In the following section, we will attempt to develop pressure test guidelines purely through analyzing reference SCG data sets. We will then review the results and check for consistency with the accelerated testing and FEM driven conclusion we reached above.

Pressure Test Guidelines

Figure 8-38 shows the process used for establishing pressure test guidelines. The basic premise is that a pressure test needs to be capable of causing damage that will breach the wall in a specified time frame under normal operating conditions to fail within the time of the over pressure test. The models needed to carry out this process are fully developed at this point. Actual guidelines will be provided in the next monthly report.
The results of steps 1, 2 and 3 are shown in **Figure 8-39** below. The data sets are for Ductile control, SCG control, LDIW control, LDIW with Indentation and LDIW with squeeze-off. There were 450 reference data points in total used in this analysis.

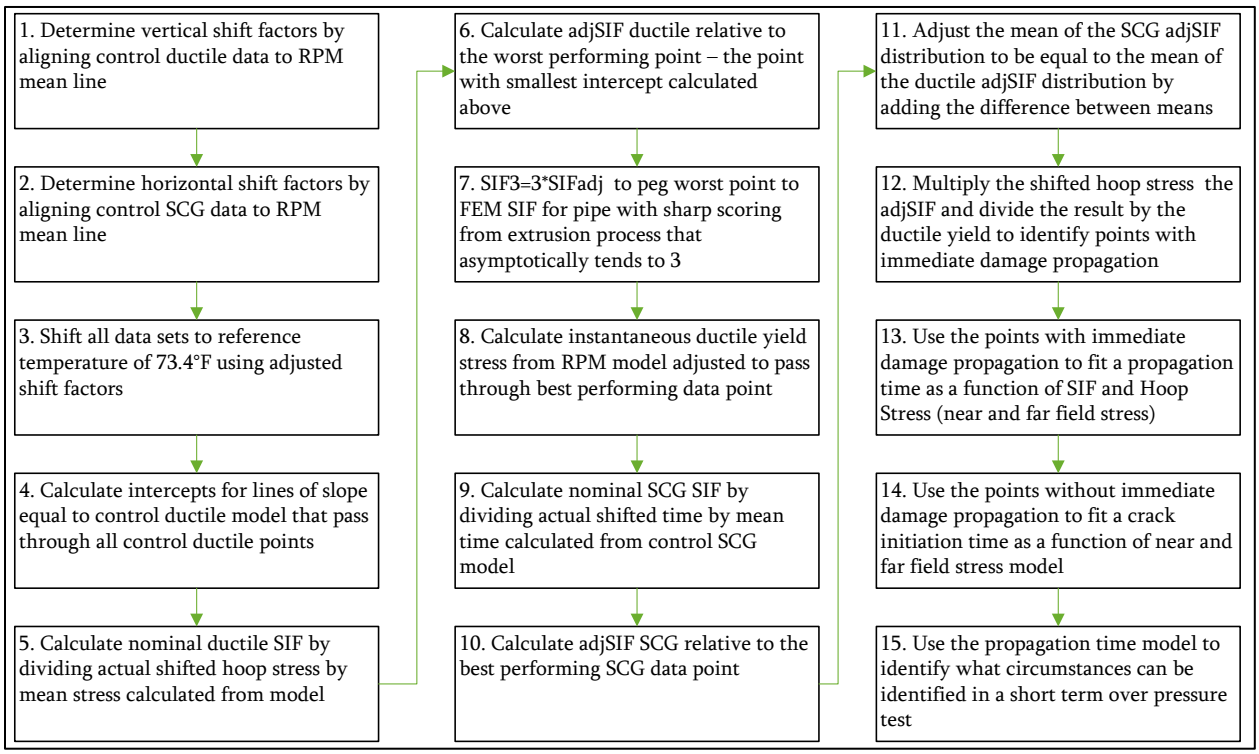


Figure 8-38. Process for establishing pressure test guidelines

FEM analyses were run to determine the SIF due to sharp and blunt grooves introduced into the pipe internal diameter during extrusion. The assumption is that these SIF "seed" the ductile rupture process and that the ductile stress rupture curves reflect these SIF. **Figure 8-40** and **Figure 8-41** show that at internal pressures typical of stress rupture tests the SIF for sharp grooves approaches 3 and for blunt grooves they approach 2.2. In steps 4, 5, 6 and 7 the distribution of the SIF in the ductile data set was calculated and is shown in **Figure 8-42.**

Fixing the maximum SIF to 3 per the FEM analysis results in a minimum SIF of 2.2, which is in remarkable agreement with the FEM analysis. This result is viewed as validation of the assumption that failure in the ductile regime is driven by the axial scoring on the pipe ID. In steps 9, 10 and 11 the SIF for the SCG data were calculated with an intermediate step of adjusting the control SCG calculated mean for the SIF distribution to match that of the ductile data points. The logic is that these are the same pipe population put into testing regimes that ensure ductile or SCG failures. The distribution of stress risers on the internal diameter of the pipes does not vary between tests. The ductile rupture process will ensure that sharp defects will blunt under plastic flow and we will see a narrow distribution of effective stress risers as reflected by the FEM results and the SIF distribution calculated from the data. In the SCG testing the far field stress is well below the ductile rupture boundary and we would expect to see a broader distribution of SIF due to less blunting of sharp notches and lower SIF also generating crack growth. This turns out to be a correct assumption. **Figure 8-43** shows the distribution of SIF in all of the reference data sets. Two parameter probability distributions have been fitted to all these results and will be used as priors in the Bayesian network models.

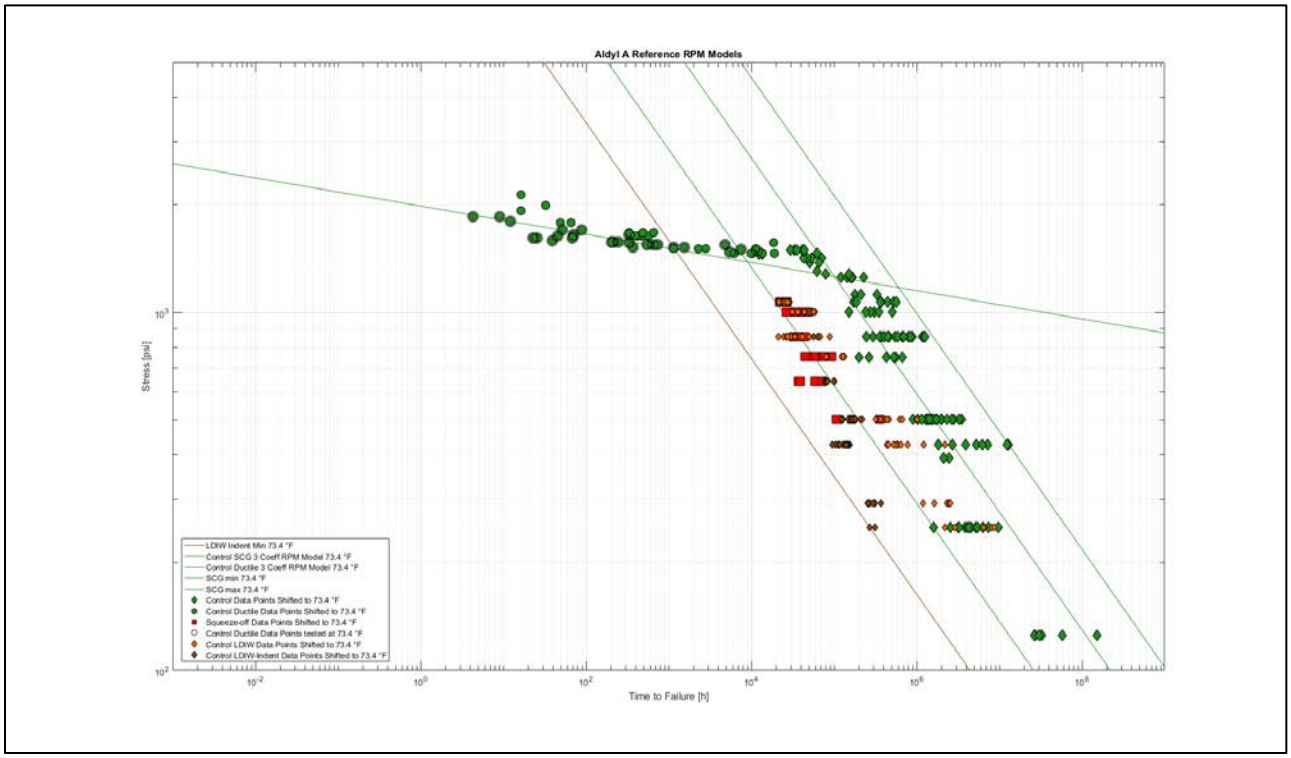


Figure 8-39. Reference data sets shifted to 73.4°F per steps 1,2 and 3 of the process for determining pressure test guidelines

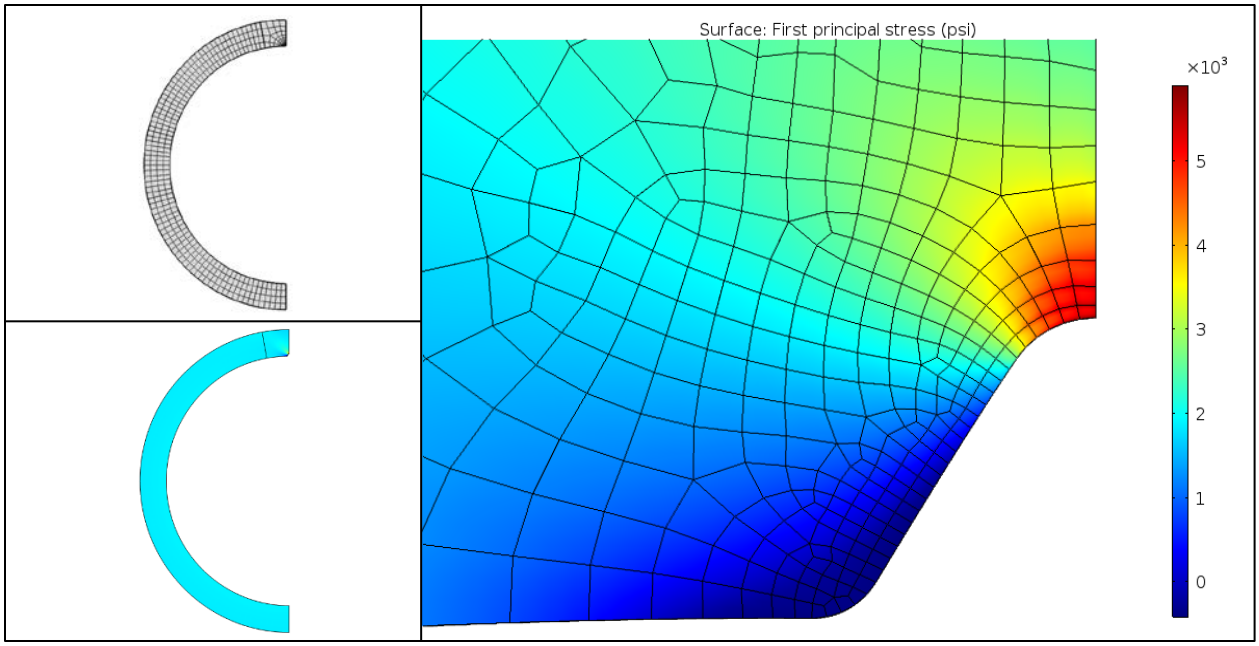


Figure 8-40. FEM analysis of stress associated with die line grooves on internal diameter of MDPE pipe

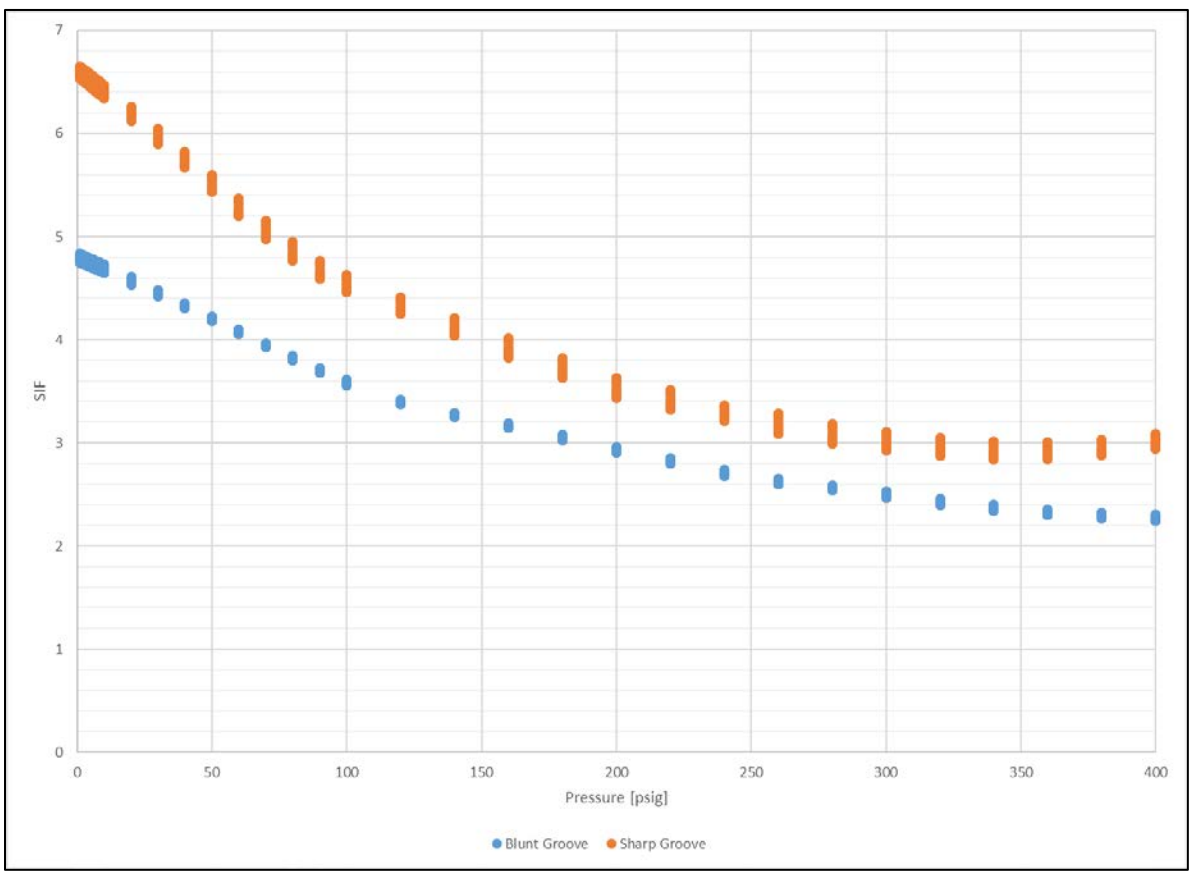


Figure 8-41. SIF measured in FEM as a function of internal pressure at 73.4°F

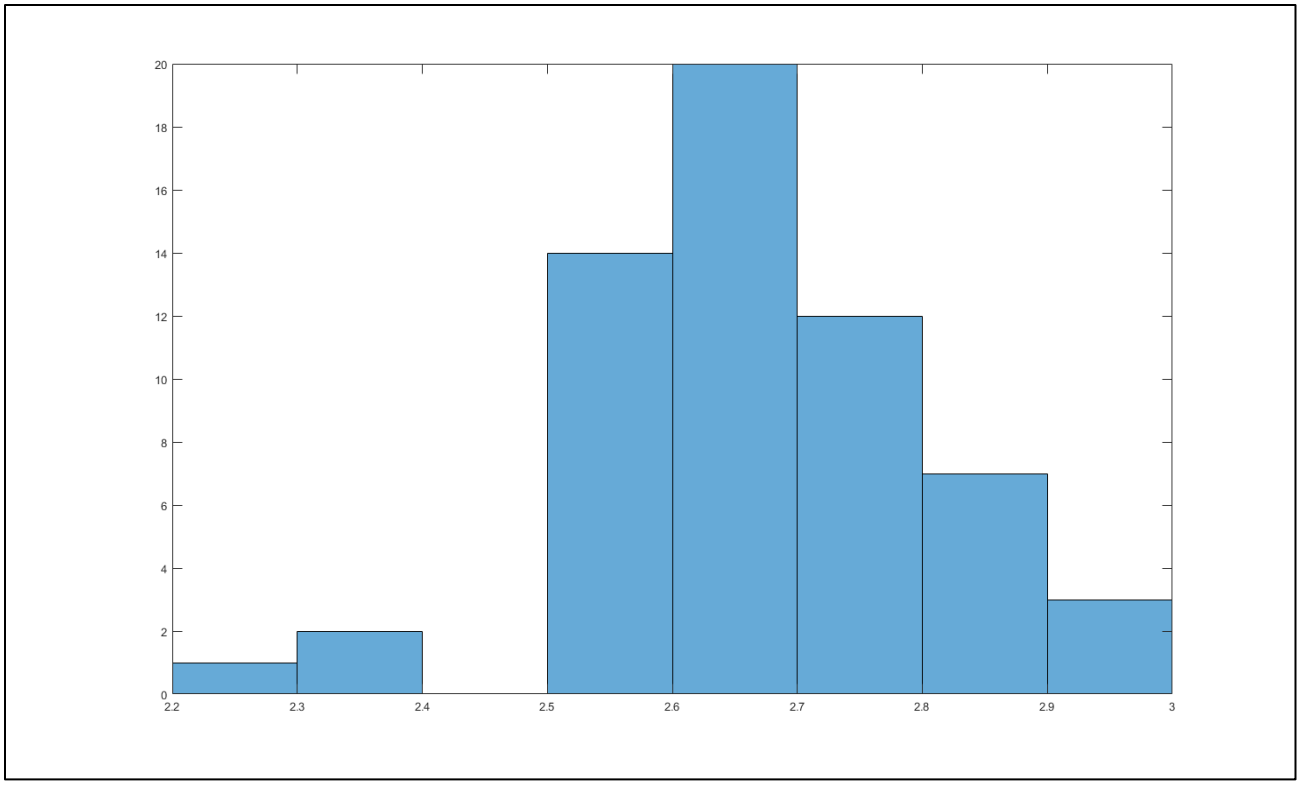


Figure 8-42. Distribution of SIF calculated in steps 4,5,6 and 7 of the process

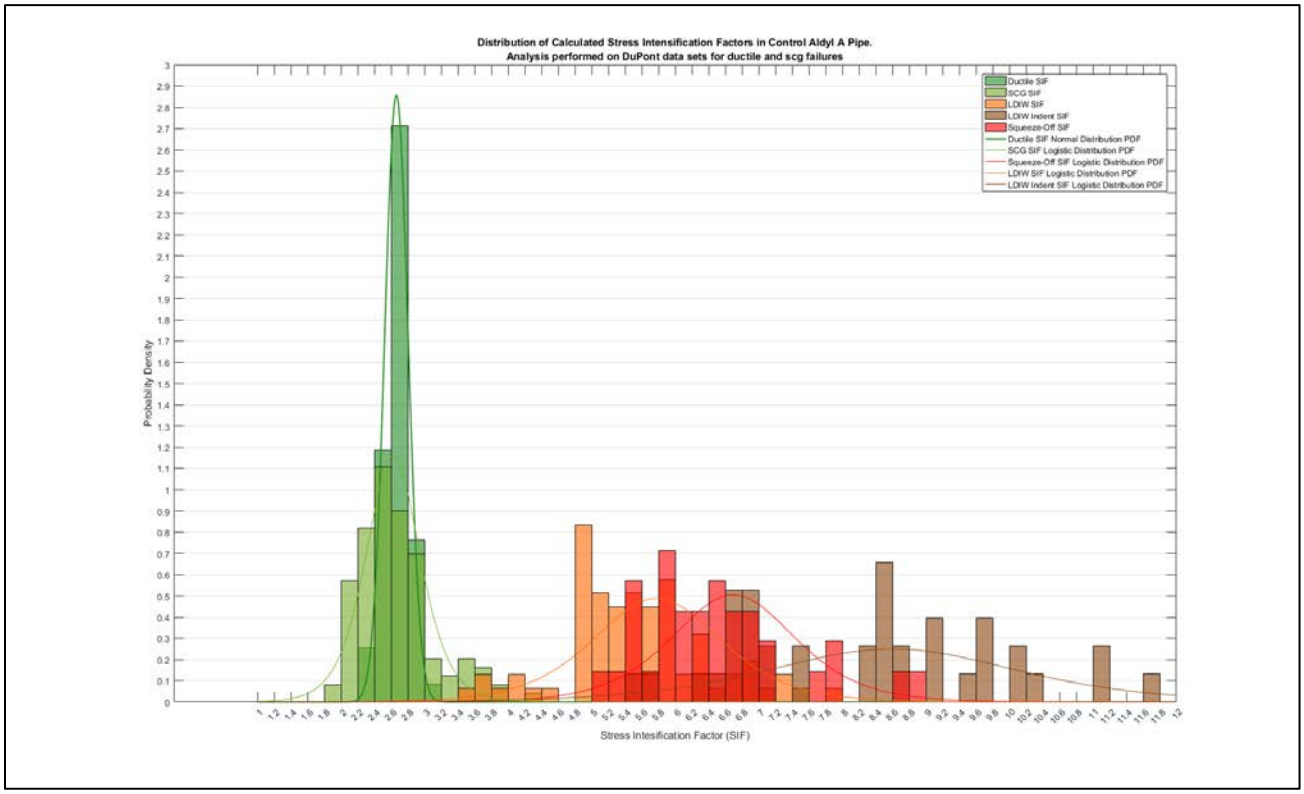


Figure 8-43. Distribution of SIF in SCG data set

The calculated SIF normalized to the mean of the ductile SIF were used to calculate the expected failure time for all data points where the nominal hoop stress multiplied by the SIF was input together with the test temperature into the control SCG RPM model. The results are shown in **Figure 8-44** with the 10 non-conservative results highlighted in red. There were 351 validation points with 2.8% of results non-conservative predictions, therefore we can conservatively state that we have 95% confidence that the model predictions will result in a conservative lifetime prediction. **Figure 8-45** shows the actual correlation estimate distribution with mean value of 0.9.

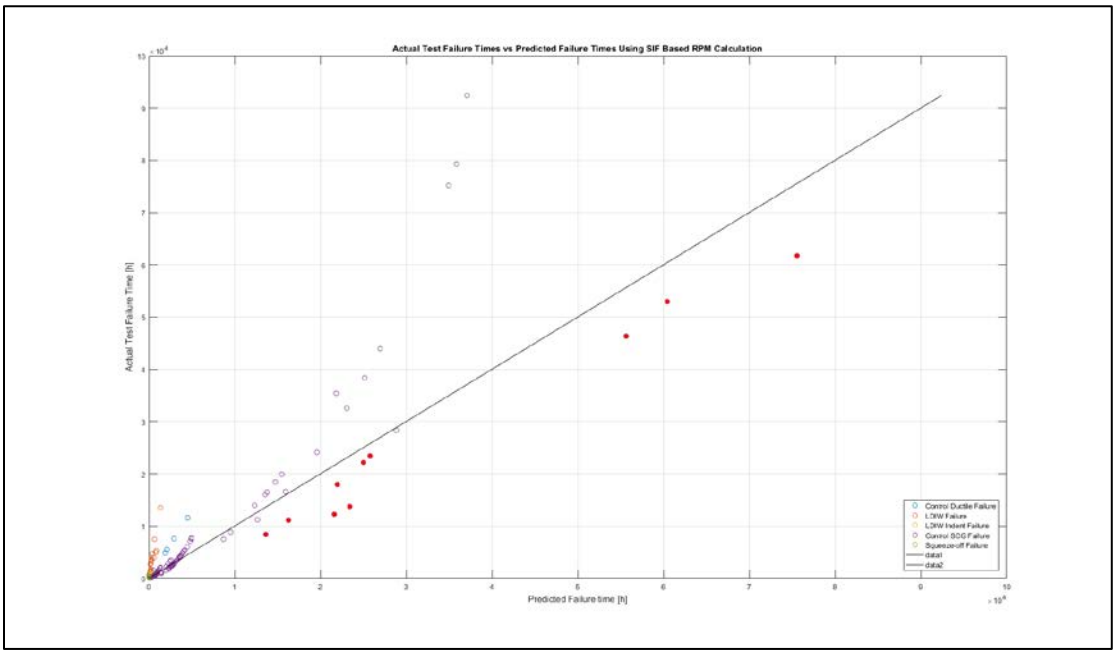


Figure 8-44. Actual failure times vs predicted failure times, red points non-conservative

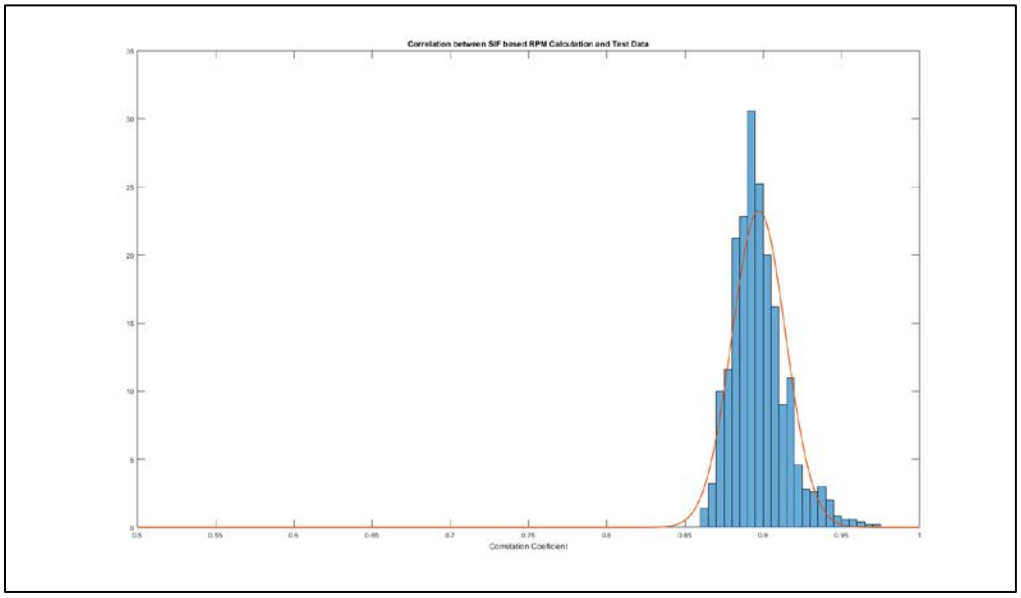


Figure 8-45. Correlation of model results to actual results mean=0.9

This result is encouraging as it validates the general approach used in the risk models employed in this project, where SIF are widely used as linear multipliers on the nominal hoop stress in conjunction with reference RPM model.

Steps 8, 12, 13 and 14 were used to extract propagation and initiation times from the data. In step 8 the instantaneous yield stress of the MDPE was determined to be 3,025 psi, which is very good agreement with mechanical testing data. This value was used to sort the SCG data into two categories:

1. Data points where the nominal hoop stress multiplied by the SIF exceeded the instantaneous yield stress, and
2. Data points where the nominal hoop stress multiplied by the SIF was less than the instantaneous yield stress.

The first group were assumed to have immediate damage propagation and the failure time was assumed to represent only propagation time. These points were used to fit a model relating the propagation time to the near field stress represented by the SIF and the far field stress represented by the nominal hoop stress. A good statistical model was found:

```
General model for Propagation Time:
    PropagationTime= a*SIF^b*HoopStress^c
Coefficients (with 95% confidence bounds):
      a =   2.233e+11 (-1.081e+11, 5.548e+11)
      b =     -1.232 (-1.459, -1.005)
      c =     -1.915 (-2.086, -1.743)
Goodness of fit:
  SSE: 5.04e+10
  R-square: 0.8644
  Adjusted R-square: 0.8628
  RMSE: 1.712e+04
```

The second group were assumed to have a crack initiation time followed by a propagation time. The propagation time for each data point was calculated from the propagation time model and subtracted from the failure time to yield the crack initiation time. The resulting data were fitted to a general model of the same form as the propagation time model:

```
General model for Initiation Time:
    InitiationTime= a*SIF^b*HoopStress^c
Coefficients (with 95% confidence bounds):
      a =   2.573e+11  (5.041e+10, 4.643e+11)
      b =     -0.8051  (-1.106, -0.5043)
      c =       -1.76  (-1.891, -1.628)
Goodness of fit:
  SSE: 4.532e+14
  R-square: 0.827
  Adjusted R-square: 0.8249
  RMSE: 1.652e+06
```

The initiation time and propagation time models were applied to all SCG data points to estimate total time to failure using hoop stress and normalized SIF as inputs.
The initiation time to total time ratio for all points was calculated. The mean was found to be 0.85, which is in good agreement with the rule of thumb that initiation time is 80-90% of the total time to failure. The distribution of results is shown in **Figure 8-46**.

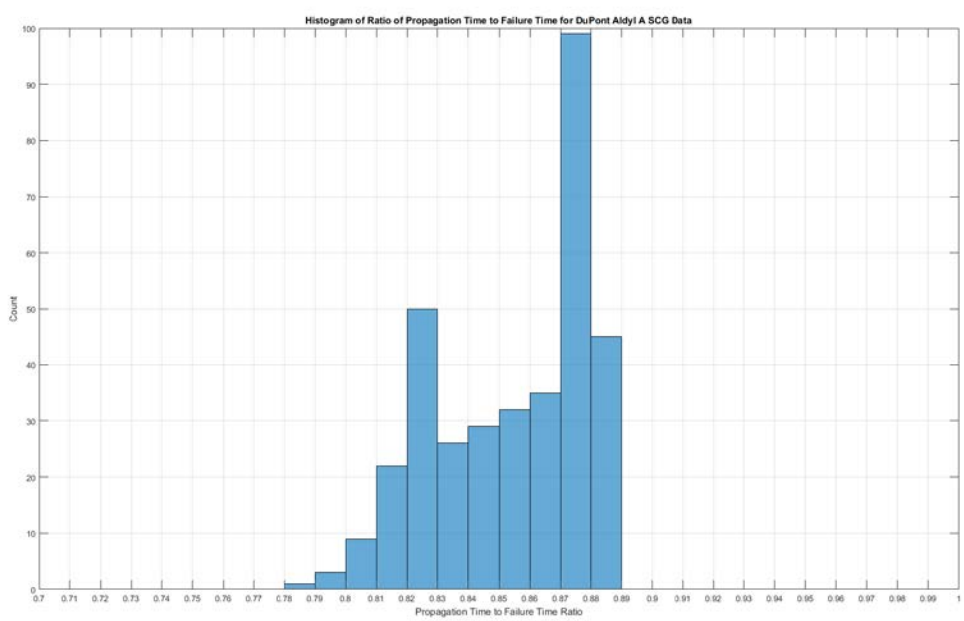


Figure 8-46. Distribution of propagation time to fail time ratio

The ratio of calculated failure time to actual failure times is shown in **Figure 8-47**.

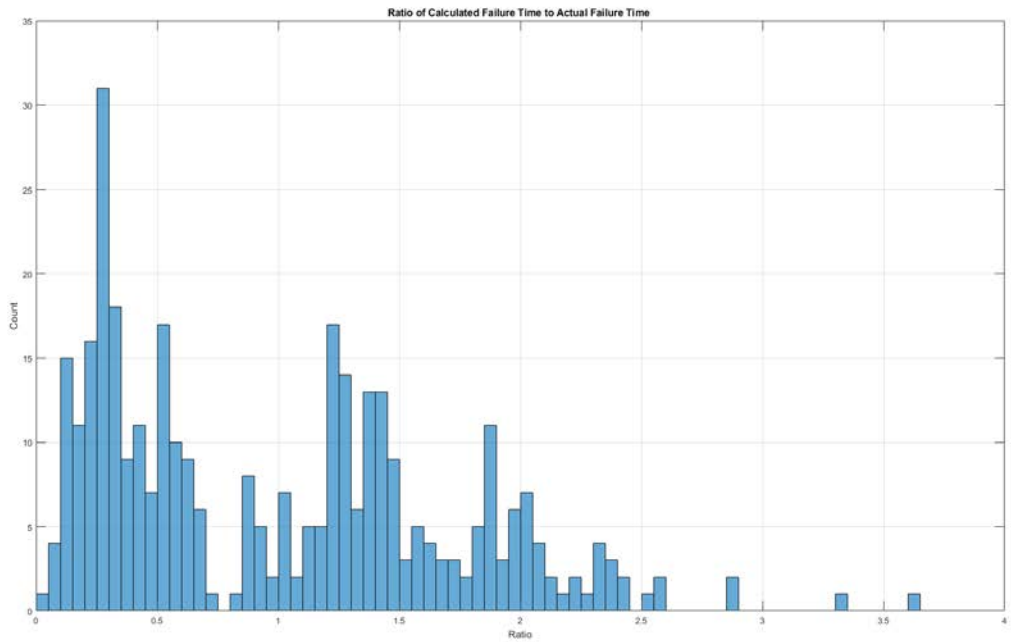


Figure 8-47. Distribution of calculated failure times to actual failure times ratio

The proper context for reviewing this result is to examine **Figure 8-39** and measure the ratio between the best performing SCG points and the worst performing points at a given hoop stress. The ratio is 50:1. If we include the problematic loading and material data this ratio grows to 250:1. For the validation shown in **Figure 8-44** all data points were included and the maximum ratio is about 3.6:1. The cumulative density plot of this distribution is shown in **Figure 8-48**. From this plot, we can see that 95% of results yield a ratio of less than 2.25 and 90% a ratio less than 2.

This result shows us that knowledge of the SIF and material condition reduces our uncertainty in lifetime prediction by up to 2 orders of magnitude.

This is a remarkable result that can be used to develop value of information statement to justify the gathering of more detailed information on the system.

Step 15, the development of pressure test guidelines is the next step in the process.

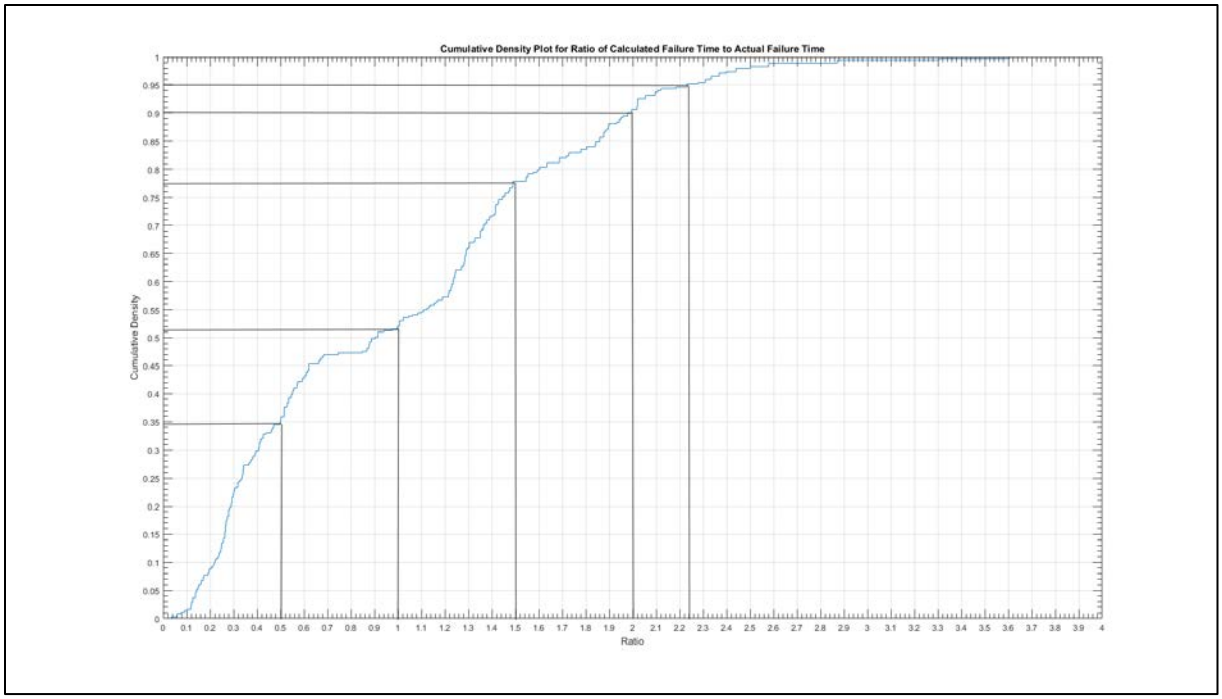


Figure 8-48. Cumulative Density Plot for ratio of calculated failure times to actual failure time

Using the model for propagation time above we can calculate the expected time to failure in an over-pressure test at 68°F where we raise the internal pressure to 1.5 * Maximum operating Pressure. The results are shown in **Figure 8-49.**

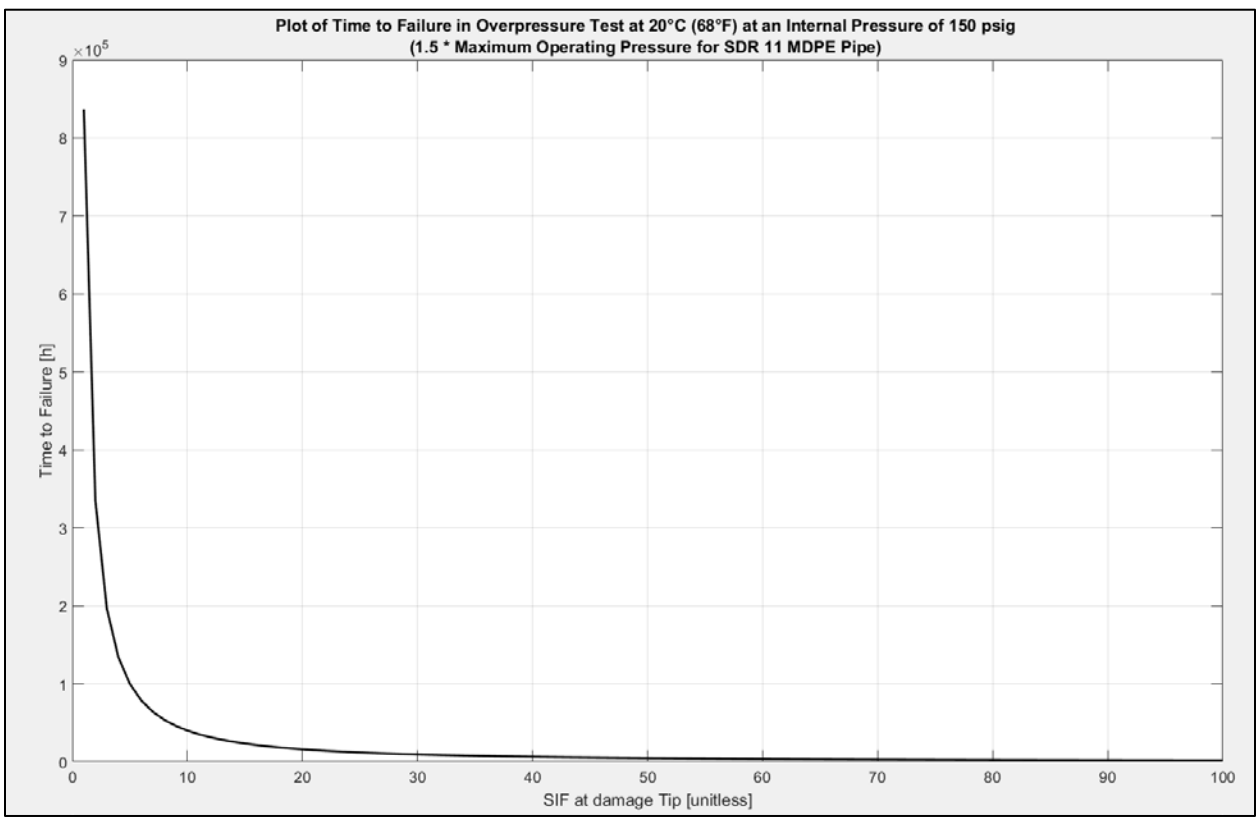


Figure 8-49. Over-pressure test @ 68°F and 150 psig internal pressure

As expected the results are absurd with failure times more than 50,000 hours at SIF <10.

Increasing the internal pressure to 400 psig per the example in the ductile tearing discussion above yields failure times in the 6,000 – 15,000-hour range for 5< SIF < 10 (**Figure 8-50**).

The maximum possible SIF for polyethylene at 400 psig internal pressure is about 11. This result is in the same ballpark as the expected time to failure derived from accelerated testing of multiple classes of polyethylene resin.

The times to failure are calibrated to the full wall thickness of SDR 11 pipe. Propagation times for fractions of the wall thickness would have to be calculated based on several assumptions about the rate of acceleration as the crack depth increases. This is beyond the scope of the present project.

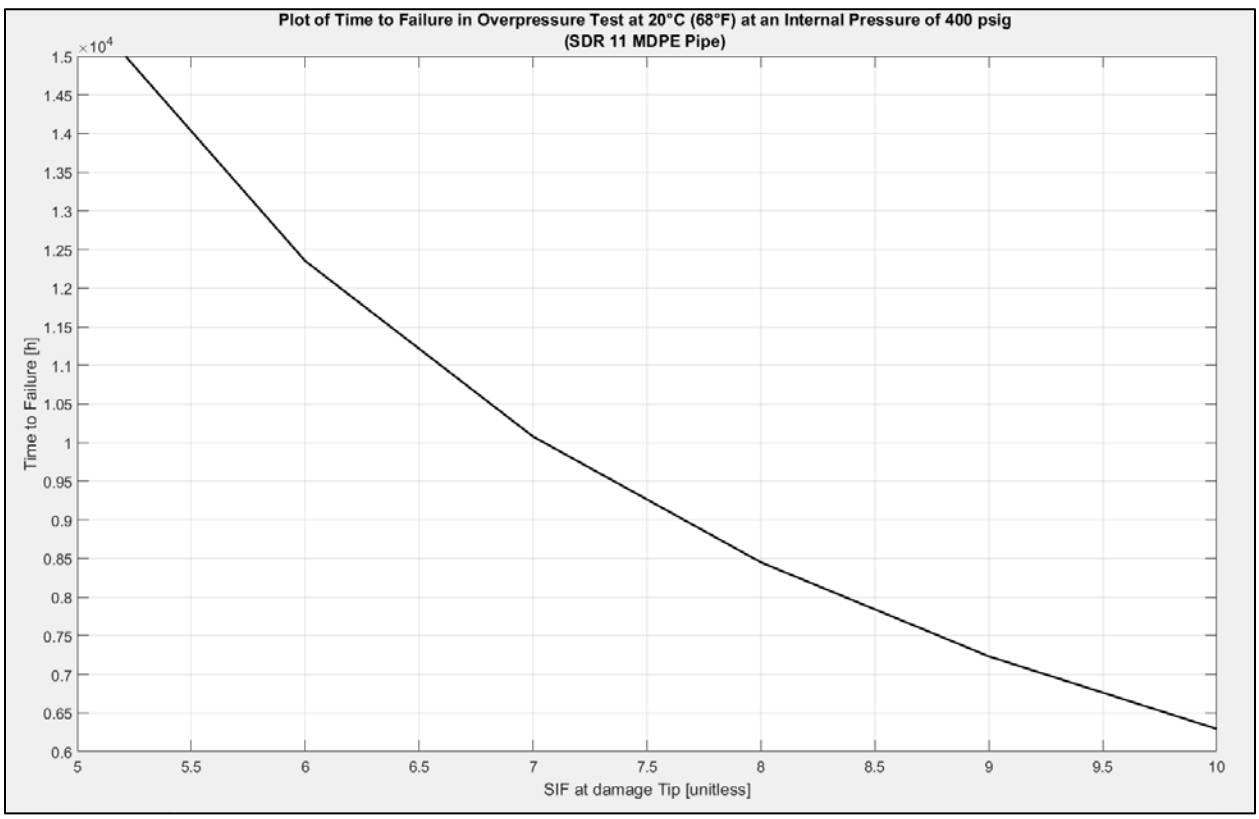


Figure 8-50. Over-pressure test @ 68°F and 400 psig internal pressure

Short-term Pressure Test Conclusion

- Polyethylene constantly creeps under load and therefore does not react in a similar manner to steel at a crack tip.
- In polyethylene damage propagates per classic ductile material models as described by Bao [2].
- At crack tip fibrils form and undergo creep at rates that depend on the local strain
- The times to failure follow a strong power law relationship to both stress and temperature
- Two approaches to estimating damage propagation times lead to similar results:
    - Propagating damage from a razor notch through 30% of the wall in accelerated testing
    - Backing propagation times out from multiple SCG reference data sets
- A short-term over pressure test will not be effective in eliminating cracks of a given size from consideration with any acceptable level of certainty.

Section 8 Bibliography

8. Lever, E., *Bi-Directional Shift Factors Revisited*, in *Plastic Pipes XVII*. 2014: Chicago, IL.
89. Kiefner, J.F. and W.A. Maxey. *The benefits and limitations of hydrostatic testing*. in *API's 51st Annual Pipeline Conference & Cybernetics Symposium, New Orleans, Louisiana (April 18-20, 2000)*. 2000.
90. Bao, Y., *Prediction of ductile crack formation in uncracked bodies*. 2003, Massachusetts Institute of Technology.
91. Bao, Y. and T. Wierzbicki, *A comparative study on various ductile crack formation criteria.* Journal of Engineering Materials and Technology, 2004. **126**(3): p. 314-324.
92. Bao, Y. and T. Wierzbicki, *On fracture locus in the equivalent strain and stress triaxiality space.* International Journal of Mechanical Sciences, 2004. **46**(1): p. 81-98.
93. Wierzbicki, T., Y. Bao, and Y. Bai. *A new experimental technique for constructing a fracture envelope of metals under multi-axial loading*. in *Proceedings of the 2005 SEM annual conference and exposition on experimental and applied mechanics*. 2005.
94. Wierzbicki, T., et al., *Calibration and evaluation of seven fracture models.* International Journal of Mechanical Sciences, 2005. **47**(4): p. 719-743.
95. Kausch, H.H., *Polymer fracture*. 2nd ed. Polymers, properties and applications. 1985, Berlin ; New York: Springer-Verlag.

9. Summary and Recommendations

Tools for Evaluating Risk

The primary objective of this project was to provide an integrated set of quantitative tools that provide a structured approach to evaluating the latent risk in vintage polyethylene pipes, such as Aldyl-A, that are common in gas distribution systems.

Polyethylene pipes undergo constant creep due to the nature of the material. The underlying molecular mechanisms that enable creep are called relaxation mechanisms and they are constantly in action. Basic molecular motions occur thousands, or millions of times per second. If there is an external driving for that loads the polymeric structure, stresses will be developed in the material. These stresses give directionality to the random molecular motions that result in creep, and ultimately lead to failure of the polyethylene structure.

A primary deliverable of this project is a set of tools that define:

4. A Rate Process Method (RPM) model that defines the rate at which the polyethylene will creep.
    a. This rate is strongly dependent on temperature
    b. The rate is also dependent on the stress in the polyethylene structure that in turn depends on:
        i. The geometry of the component
        ii. The external loads acting on the component
5. The Stress Intensification Factors (SIF) that provide a simple means of translating the nominal hoop stress of the pipe, which is very easy to calculate, to true stress. Well defined and simple to apply SIF are essential to a workable risk evaluation method that utilizes simple, well-known parameters such as: pipe size, ambient temperature, system operating pressure, component configuration and other measurable installation characteristics to arrive at a true component stress. The SIF developed in this project, together with a single master RPM model underpin the lifetime prediction methods presented in this project.
6. The RPM model and SIF can be used as-is to perform risk assessments given system parameters, or they can be integrated into a tool that is capable of integrating all threat interactions into a composite risk score. A deliverable of this project is a Bayesian network that accomplishes this objective. A fully defined, calibrated and validated Bayesian network is defined in this project.

Non-Destructive Evaluation in Confined Spaces, Fitness for Service, Replacement Prioritization and Data integration

Secondary objectives of the project were: first, to provide a fitness for service approach that can support replacement prioritization; second, utilize data from multiple sources such as in ditch condition assessment and leak records; third, to provide a means to access the pipe in a congested urban environment. These objectives were technically realized through:

4. A non-destructive tool that is capable of measuring pipeline configuration from inside the pipe was developed and prototyped in this project. This structured light, endoscopic measurement tool is a major breakthrough in assessing gas distribution pipes as it allows the operator to measure pipeline geometry over large lengths of the pipe without excavating the entire pipe. This measurement of pipe geometry from inside the pipe allows identification of several critical defects such as: impingement, squeeze-off, fittings, sudden displacements of the pipe, pipe deformations and other defects that cause stress intensification. This direct measurement of features will allow accurate SIF to be assigned to pipe segments. This will allow proper classification of segment with regard to the anticipated stress fields that when plugged in to the RPM model will provide probability of failure over time. This likelihood of failure is a key component in determining the segments Fitness for Service (FFS).
5. A set of reliability based tools were developed that underpin optimization methods for comparing repair/replace strategies over multi-year timeframes. These methods are based on robust damage propagation methods that were calibrated and validated against historic reference data. It was demonstrated that the Monte Carlo simulations that these tool support are capable of evaluated multiple scenarios and providing guidance as to the most effective risk management strategy over time
6. The tools developed in this project were integrated into a commercially available Artificial Intelligence (AI) platform that is capable of merging multiple disparate data sources, running the various risk assessment tools and providing insights driven be sophisticated data analytics. Intelligent forms that facilitate in-field data gathering and regulatory reporting requirement were also developed and demonstrated and tested as part of this project

Summary

All of the project deliverables were met and tested via the components described above. A comprehensive set of tools that can be practically applied in multiple approaches, from simpe point applications, to enterprise wide decision support systems has been provided.

Recommendations

The RPM and SIF based methods developed in this project can be immediately applied in:

1. Relative risk models
2. Quantitative risk models
3. Probabilistic risk models

They can be applied to point problems, as well as system wide problems. However, to address system wide problems in a coherent manner it is advisable to transition to probabilistic decision support systems. This is not a simple task. A Joint Industry Program (JIP) involving all stakeholders i.e. regulators, operators, material suppliers, manufacturers should be established to develop a roadmap that provides guidelines to the industry on how to improve all risk assessment methods outlined above through adoption of the tools provided by this project. The roadmap should include transition pathways between each of the risk assessment methodologies.

A key first effort should be to provide clear guidelines on how to improve data collection methods to support probabilistic reasoning and reliability based risk assessment methods. Efficient data collection will be greatly enhanced by the structured light scanning method developed in this project. The JIP should be tasked with shepherding the commercial development and deployment of this exciting technology that can be integrated into existing keyhole technologies.

The JIP can provide useful guidance on how to use the methods developed inn this project to support transitioning from prescriptive regulation to performance based regulation that will have reliability based maintenance frameworks as a key underpinning.

10. Future Work

The results presented in this report are a synthesis of five discreet efforts performed by three institutions GTI, ASU and MSU. GTI defined the architecture of the project and supervised the collaborative effort. GTI developed the adjusted RPM model and SIF distributions essential for performing fitness for service calculations, a Bayesian network for synthesizing complete system information, and integrated the above into a commercially available AI system together with developing an intelligent field data collection framework. MSU developed the structured light damage detection prototypes and associated algorithms. ASU developed methods for recognizing damage from the structured light data stream, and developed a maintenance optimization framework based on damage propagation and reliability methods that can be integrated into an enterprise decision support framework.

At the end of this collaborative effort considerable thought was put into identifying future research efforts that will enhance the methods presented. Some of the suggestions are presented below for consideration by interested parties.

Fitness for Service

The presented work provides an excellent baseline that will enable operators to immediately apply the methods to point problems as they are identified in their systems. Several method improvements are needed to facilitate the transition to probabilistic risk assessment in the context of enterprise decision support systems:

1. Work needs to be done to extend the concepts presented into simulation the SCG process in polyethylene materials.
    a) The strain dependency of bidirectional shift factors needs to be investigated in the strain hardening portion of the stress/strain curve
    b) Creep rate experiments need to be conducted in the strain hardening portion of the stress/strain curve
    c) Bao-Wierzbicki models need to be developed to describe the critical strain at break of polyethylene fibrils
2. Implementation projects need to be carried out to develop the intelligent field data collection concepts presented in this project. These projects will develop the sampling plans and data categories needed to support probabilistic reasoning in system wide fitness for service determinations.
3. Work should be performed to integrate the decision support methods outlined in this project into enterprise risk governance frameworks i.e. risk tolerance and objective setting frameworks need to be developed to facilitate progressing beyond compliance towards true risk management that can inform the risk governance efforts of the organization.

Decision Support systems

The presented work meets the goal in the proposal statement. Some additional work can be done to enhance and further develop the maintenance framework presented in this study:

1. The algorithm for the image reconstruction can be adjusted. The camera needs to be calibrated so that it can calculate the coordinate of the patterned light relative to the camera. With a calibrated camera, the size of the damage can also be known which is useful in the proposed dynamic maintenance framework. The dimensional information could be used to update the stress concentration factor and hence update the maintenance plan.
2. Although the proposed creep prediction model agrees well with the experimental data, there still exists some bias from the lab data. Further demonstration is needed if there were additional data available from GTI.
3. The ductile group in the data set was not used, but a trend of smooth transition can be observed in the data point plot (Figure 6-15) from ductile data to SCG data. An asymptotic function regarding yield strength of the material can be modeled for the slope of the regression curve. This would need further research and justification.
4. The maintenance framework is flexible in solving the maintenance planning given different conditions (minimize cost given reliability constraint and optimize reliability given budget constraint). The computational cost would be enormous when applied in large scale systems with various groups of pipes. The underlying principle when dealing with large scale optimization is to be studied. The maintenance framework may need to be adjusted for the high dimensional problem.
5. The proposed Bayesian entropy network (BEN) as a classifier can achieve fast learning with the extra information. The BEN concept can also be used in updating probabilities. The first step is applying the network in the updating of the parameters in the creep crack prediction model. The BEN can also be applied into large scale systems. In a large network, the update of one parameter would affect the nodes in the whole system. Sometimes this effect is not wanted. With the additional constraint, the BEN updating can be more accurate and efficient than traditional Bayesian network updating.
6. Sometimes, there could be a bias between the expert opinion (empirical information) and the ground truth. Once the wrong information were coded into the network, instead of increasing the inference accuracy, it would degrade the performance of the network and even damage the whole system. Since the constraint in a BEN framework is strong, an adaptive BEN is needed to compensate this situation. The adaptive network would use the given constraint when there is not enough data to update the belief. But would change to the truth from data when more observation become available. This could be understood as when there is limited information, the

network chooses to believe experience, but shifts its belief to the truth brought by data when more evidence become available.

7. The current framework focuses on one failure mode of plastic pipelines (slow crack growth). Information fusion and big data analytics with multiple failure modes and large systems that can integrate with the ongoing development of GTI framework needs further development. The "agent" serving for different failure types will need to automatically fused together for a consistent risk assessment. Recent advancement in artificial intelligence (AI) such as deep network learning has the potential for diagnostics and prognostics for gas pipeline industries.

Sensing Damage via Internal Inspection

The delivered work meets the goal and proposed research objectives. Some additional research and hardware development can be done to enhance the ESLiST sensing system:

1. The current system provides a high quality and fast surface profile reconstruction, but the spatial resolution can be further increased by improving the hardware design, better system calibration, and more efficient reconstruction algorithm. Currently, a fixed structured light pattern was applied, which might not be optimal for all damage types and different field testing conditions. With the successful demonstration of the Prototype, I to Prototype IV and their feasibility of extracting damage dimensional information at MSU, extensive involvement of GTI in collecting real data should be expected. Additional work on feeding updated stress concentration factor derived from ESLiST image and ESLiST assisted decision information back to the sensing system should be carried out to complete the development loop and further optimize the novel structured light sensors design.
2. The proposed prototype works very well in a controlled lab environment by assuming the sensor is moving in a controlled straight path that is along the pipe's center axis. This assumption might be violated in actual field testing and inspection due to mechanical vibration of the sensing platform moving inside the pipe with complex geometries. An inertial measurement unit (IMU) can be integrated into the system to provide an estimation of the device orientation and positioning. Real-time compensation algorithms are needed to correct any distortion and misalignment induced error in data reconstruction to achieve same imaging resolution in a field environment.
3. Static structured light patterns have been studied and implemented as we proposed. However, the potential of the structured light approach will be maximized by introducing dynamic light patterns and data structures to adapt to various damage types and sizes under different light conditions. This could be better understood through both numerical simulations and experimental studies by collaborating with GTI and industry partners. An optimization of the current static structure light patterns to study the

relationship between the smallest detectable damage feature and the parameters of the static light pattern will also be beneficial.

4. Further improvement of the current data reduction and reconstruction will lead to faster and more accurate damage detection. The current framework only focuses on the SCG failure mode, so supervised multispectral data dimensionality reduction and defect classification methods would be sufficient and are successfully demonstrated. However, additional work should be done to realize the unsupervised damage recognition and information fusion while multiple failure modes are considered. Integration of advanced sensors from MSU and the intelligent decision support system from GTI and ASU would be more critical to achieving an optimal diagnostic and prognostic framework for gas pipeline industries.

List of Acronyms

Acronym	Description
ALM	Augmented LaGrange multiplier
BEN	Bayesian/maximum Entropy Network
CAD	Computer aided design
CBM	Condition-Based Maintenance
CCD	Charge coupled device
CCG	Creep Crack Growth
CDF	Cumulative Density Function
CNN	Convolutional neural network
CTS	Copper tube size
DLP	Digital light processing
DOF	Degree of Freedom
EDSS	Enterprise Decision Support System
ESLIST	Endoscopic structured light scanning tool
FEM	Finite Element Method
FFS	Fitness for Service
FOV	Field of view
GA	Genetic Algorithm
GTI	Gas Technology Institute
HSV	Hue, Saturation, Value
ILI	In-line inspection
IMU	Inertial measurement unit
LDIW	Low Ductile Inner Wall
LEAP	Laboratory of Electromagnetic and Acoustic Imaging and Prognostics
LED	Light emitting diode
LIDAR	Light detection and ranging
MC	Monte Carlo
MCMC	Markov Chain Monte Carlo
ME	Maximum Entropy
NB	Naïve Bayes

Acronym	Description
NDE	Nondestructive evaluation
NN	Neural Network
PDF	Probability Density Function
PI	Project investigator
POD	Probability of detection
POD	Probability of detection
PVC	Polyvinyl Chloride
RGB	Red green blue
RPM	Rate Process Model/Method
RUL	Remaining Useful Life
SCG	Slow crack growth
SCG	Slow crack growth
SEM	Scanning Electron Microscope
SIF	Stress intensity factor
SS	Single shot
TAN	Tree Augmented Naïve Bayes

References

1. Kurdziel, J.M. and E.F. Palermo. *Stress Crack Resistance of Structural Members in Corrugated High-Density Polyethylene Pipe*. in *Transportation Research Board 86th Annual Meeting*. 2007.
2. ISO, *9080:2012 Plastics piping and ducting systems -- Determination of the long-term hydrostatic strength of thermoplastics materials in pipe form by extrapolation*. 2012, ISO.
3. Bragaw, C. *Service rating of polyethylene piping systems by the rate process method*. in *Proceedings Eighth Plastic Fuel Gas Pipe Symposium*. 1983.
4. Bragaw, C.G. *The Forecast of Polyethylene Pipe and Fitting Burst Life Using Rate Process Theory*. in *5th International Conference on Plastic Pipes, York UK*. 1982.
5. Coleman, B.D., *Application of the theory of absolute reaction rates to the creep failure of polymeric filaments.* Journal of Polymer Science, 1956. **20**(96): p. 447-455.
6. Montgomery, D.C. and G.C. Runger, *Applied statistics and probability for engineers*. 4th ed. 2007, Hoboken, NJ: Wiley. xvi, 768 p.
7. Lever, E., *Remaining Life Expectancy and Risk Profile of Vintage Aldyl A Piping*, in *Plastic Pipes XVII*. 2014: Chicago.
8. Lever, E., *Bi-Directional Shift Factors Revisited*, in *Plastic Pipes XVII*. 2014: Chicago, IL.
9. Mavridis, H. and R. Shroff, *Temperature dependence of polyolefin melt rheology.* Polymer Engineering & Science, 1992. **32**(23): p. 1778-1791.
10. Popelar, C. *A Comparison of the Rate Process Method and the Bidirectional Shifting Method*. in *Proceedings of the Thirteenth International Plastic Fuel Gas Pipe Symposium*. 1993.
11. Popelar, C., C. Popelar, and V. Kenner, *Viscoelastic material characterization and modeling for polyethylene.* Polymer Engineering & Science, 1990. **30**(10): p. 577-586.
12. Strobl, G.R., *The Physics of Polymers: Concepts for Understanding Their Structures and Behavior*. 2013: Springer Berlin Heidelberg.
13. Bower, D.I., *An Introduction to Polymer Physics*. 2002: Cambridge University Press.
14. Popelar, C.F., *Characterization of mechanical properties for polyethylene gas pipe materials*. 1989, The Ohio State University.
15. Frank, A., G. Pinter, and R.W. Lang, *Prediction of the remaining lifetime of polyethylene pipes after up to 30 years in use.* Polymer Testing¶, 2009. **28¶**(7¶): p. 737-745.
16. Young, W.C. and R.G. Budynas, *Roark's Formulas for Stress and Strain*. 2001: McGraw-Hill Education.
17. Becht, C., *Process Piping: The Complete Guide to ASME B31.3*. 2009: ASME Press.
18. Maupin, J.M., Michael, *DTPH56-06-T-0004: Plastic Pipe Failure, Risk,and Threat Analysis*. 2009, GTI.
19. Pearl, J., *Causality: Models, Reasoning, and Inference* 2000: Cambridge University Press.
20. Fenton, N. and M. Neil, *Risk assessment and decision analysis with Bayesian networks*. 2012: CRC Press.
21. Haine, S., *Hazard Analysis & Mitigation Report on Aldyl A Polyethylene Gas Pipelines in California*. 2014, California Public Utilities Commission.
22. Bianca Pinheiro, I.P.a.S.B.C. *Stress Concentration Factors of Dented Pipelines*. in *6th International Pipeline Conference*. 2006. Calgary.

23. Keating, A.J.R.a.P.B., *Stress Concentration Solution for a 2D Dent in an Internally Pressurized Cylinder.* JOURNAL OF ENGINEERING MECHANICS, 2007. **133**(7): p. 792-800.
24. Godoy, L.A., *Thin-walled structures with structural imperfections*. 1996, New York: Pergamon.
25. Jajo, J.O., *Dent behaviour of steel pipes under pressure load*. 2014, Electronic Theses and Dissertations.
26. *Multinomial logistic regression*, in *Wikipedia*.
27. Boenisch, A. *Advanced subsea pipeline inspection: an overview of flied experience and benefits of subsea pipeline*. in *Pipelines International*. 2013.
28. Barnwell, C., W. Doyle, and C. Gunn, *LiDAR for terrain mapping on the Alaska pipeline corridor.* Office of Federal Coordinator, accessed at http://www. arlis. org/docs/vol1/AlaskaGas/Paper, 2009.
29. Duran, O., K. Althoefer, and L.D. Seneviratne, *Automated pipe defect detection and categorization using camera/laser-based profiler and artificial neural network.* IEEE Transactions on Automation Science and Engineering, 2007. **4**(1): p. 118-126.
30. Zhang, L., B. Curless, and S.M. Seitz. *Rapid shape acquisition using color structured light and multi-pass dynamic programming*. in *3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on*. 2002. IEEE.
31. Drouin, M.-A. and J.-A. Beraldin, *Active 3D Imaging Systems*, in *3D Imaging, Analysis and Applications*. 2012, Springer. p. 95-138.
32. *CCTV Pipeline Testing Survey & Reporting with Wincan v8*. 2013]; Available from: http://www.auscamqld.com.au/.
33. Agency, I.A.E., *Training Guidelines In Non-Destructive Testing Techniques*. 2013, International Atomic Energy Agency.
34. Newsletter, S.C.G. *Applications of Computer Vision to Computer Graphics*. 1999 November 1999]; Available from: https://www.siggraph.org/publications/newsletter/v33n4/contributions/curless.html.
35. Gonzalez, R.C., S. Eddins, and R.E. Woods, *Digital image processing using MATLAB*. 2004.
36. Salvi, J., et al., *A state of the art in structured light patterns for surface profilometry.* Pattern recognition, 2010. **43**(8): p. 2666-2680.
37. Geng, J., *Structured-light 3D surface imaging: a tutorial.* Advances in Optics and Photonics, 2011. **3**(2): p. 128-160.
38. Monks, T. and J.N. Carter. *Improved stripe matching for colour encoded structured light*. in *International Conference on Computer Analysis of Images and Patterns*. 1993. Springer.
39. Boyer, K.L. and A.C. Kak, *Color-encoded structured light for rapid active ranging.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987(1): p. 14-28.
40. Guan, C., L. Hassebrook, and D. Lau, *Composite structured light pattern for three-dimensional video.* Optics Express, 2003. **11**(5): p. 406-417.
41. Schmalz, C. and E. Angelopoulou. *A graph-based approach for robust single-shot structured light*. in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. 2010. IEEE.
42. Schmalz, C., et al., *An endoscopic 3D scanner based on structured light.* Medical Image Analysis, 2012. **16**(5): p. 1063-1072.

43. Commons, W. *Ellipse Properties of Directrix.svg*. Available from: https://commons.wikimedia.org/wiki/File:Ellipse_Properties_of_Directrix.svg.
44. Pi, R. *Raspberry Pi 2 Model B*. Available from: https://www.raspberrypi.org/products/raspberry-pi-2-model-b/.
45. Young, M., et al. *coded structured Light*. in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. 2007. IEEE.
46. Turk, G. and M. Levoy, *The stanford bunny*. 2005, Stanford University Computer Graphics Laboratory, URL http://graphics.stanford.edu/data/3Dscanrep.
47. Je, C., S.W. Lee, and R.-H. Park. *High-contrast color-stripe pattern for rapid structured-light range imaging*. in *European Conference on Computer Vision*. 2004. Springer.
48. Zhang, Z., *A flexible new technique for camera calibration.* IEEE Transactions on pattern analysis and machine intelligence, 2000. **22**(11): p. 1330-1334.
49. Lin, Z., M. Chen, and Y. Ma, *The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices.* arXiv preprint arXiv:1009.5055, 2010.
50. O'Hagan, A. and M.G. Kendall, *Kendall's Advanced Theory of Statistics: Bayesian inference. Volume 2B*. 1994: Edward Arnold.
51. Bayes, M. and M. Price, *An essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, frs communicated by mr. price, in a letter to john canton, amfrs.* Philosophical Transactions (1683-1775), 1763: p. 370-418.
52. Jayech, K., M.A. Mahjoub, and N. Ghanmi. *Application of Bayesian networks for pattern recognition: Character recognition case*. in *Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), 2012 6th International Conference on*. 2012. IEEE.
53. Likforman-Sulem, L. and M. Sigelle, *Recognition of degraded characters using dynamic Bayesian networks.* Pattern Recognition, 2008. **41**(10): p. 3092-3103.
54. Jayech, K. and M.A. Mahjoub, *Clustering and Bayesian network for image of faces classification.* arXiv preprint arXiv:1204.1679, 2012.
55. Zhang, L. and Q. Ji, *A Bayesian network model for automatic and interactive image segmentation.* IEEE Transactions on Image Processing, 2011. **20**(9): p. 2582-2593.
56. Friedman, N., D. Geiger, and M. Goldszmidt, *Bayesian network classifiers.* Machine learning, 1997. **29**(2-3): p. 131-163.
57. Langley, P. and S. Sage. *Induction of selective Bayesian classifiers*. in *Proceedings of the Tenth international conference on Uncertainty in artificial intelligence*. 1994. Morgan Kaufmann Publishers Inc.
58. Kohavi, R. *Scaling Up the Accuracy of Naive-Bayes Classifiers: A Decision-Tree Hybrid*. in *KDD*. 1996.
59. Jaynes, E.T., *Information theory and statistical mechanics.* Physical review, 1957. **106**(4): p. 620.
60. Jaynes, E.T., *Information theory and statistical mechanics. II.* Physical review, 1957. **108**(2): p. 171.
61. Caticha, A. and A. Giffin. *Updating probabilities*. in *AIP Conference Proceedings*. 2006. AIP.
62. Landes, J. and J. Begley, *A fracture mechanics approach to creep crack growth*, in *Mechanics of crack growth*. 1976, ASTM International.

63. Williams, J. and G. Marshall. *Environmental crack and craze growth phenomena in polymers*. in *Proceedings of the royal society of London A: mathematical, physical and engineering sciences*. 1975. The Royal Society.
64. Kramer, E.J. and E.W. Hart, *Theory of slow, steady state crack growth in polymer glasses.* Polymer, 1984. **25**(11): p. 1667-1678.
65. Newman, J. and I. Raju, *Stress intensity factors for internal surface cracks in cylindrical pressure vessels.* Journal of Pressure Vessel Technology, 1980. **102**(4): p. 342-346.
66. Liu, Y. and S. Mahadevan, *Fatigue limit prediction of notched components using short crack growth theory and an asymptotic interpolation method.* Engineering Fracture Mechanics, 2009. **76**(15): p. 2317-2331.
67. Jardine, A.K., D. Lin, and D. Banjevic, *A review on machinery diagnostics and prognostics implementing condition-based maintenance.* Mechanical systems and signal processing, 2006. **20**(7): p. 1483-1510.
68. Ma, J. and J.C. Li, *Detection of localised defects in rolling element bearings via composite hypothesis test.* Mechanical Systems and Signal Processing, 1995. **9**(1): p. 63-75.
69. Davies, C. and R. Greenough. *The use of information systems in fault diagnosis*. in *ADVANCES IN MANUFACTURING TECHNOLOGY-CONFERENCE-*. 2000. TAYLOR & FRANCIS LTD.
70. Siddique, A., G. Yadava, and B. Singh. *Applications of artificial intelligence techniques for induction machine stator fault diagnosis*. in *Proceedings of 4 IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics, and Drives*. 2003.
71. Loparo, K.A., et al., *Fault detection and diagnosis of rotating machinery.* IEEE Transactions on Industrial Electronics, 2000. **47**(5): p. 1005-1014.
72. Howard, I., S. Jia, and J. Wang, *The dynamic modelling of a spur gear in mesh including friction and a crack.* Mechanical systems and signal processing, 2001. **15**(5): p. 831-853.
73. Banjevic, D. and A. Jardine, *Calculation of reliability function and remaining useful life for a Markov failure time process.* IMA journal of management mathematics, 2006. **17**(2): p. 115-130.
74. Vlok, P.-J., M. Wnek, and M. Zygmunt, *Utilising statistical residual life estimates of bearings to quantify the influence of preventive maintenance actions.* Mechanical systems and signal processing, 2004. **18**(4): p. 833-847.
75. Kwan, C., et al. *A novel approach to fault diagnostics and prognostics*. in *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on*. 2003. IEEE.
76. Yam, R., et al., *Intelligent predictive decision support system for condition-based maintenance.* The International Journal of Advanced Manufacturing Technology, 2001. **17**(5): p. 383-391.
77. Guan, X., R. Jha, and Y. Liu, *Model selection, updating, and averaging for probabilistic fatigue damage prognosis.* Structural Safety, 2011. **33**(3): p. 242-249.
78. Wang, W.-b., *A model to determine the optimal critical level and the monitoring intervals in condition-based maintenance.* International Journal of Production Research, 2000. **38**(6): p. 1425-1436.
79. Peng, T., et al., *Integrated fatigue damage diagnosis and prognosis under uncertainties*. 2012, NATIONAL AERONAUTICS AND SPACE ADMINISTRATION MOFFETT FIELD CA AMES RESEARCH CENTER.

80. Peng, T., et al., *Probabilistic fatigue damage prognosis of lap joint using Bayesian updating.* Journal of Intelligent Material Systems and Structures, 2015. **26**(8): p. 965-979.
81. Gould, R.G., et al. *Investigation of a video frame averaging digital subtraction fluoroscopic system*. in *Proc. SPIE*. 1981.
82. Elgammal, A., D. Harwood, and L. Davis, *Non-parametric model for background subtraction.* Computer Vision—ECCV 2000, 2000: p. 751-767.
83. Demuth, H. and M. Beale, *Neural Network Toolbox For Use with Matlab--User'S Guide Verion 3.0.* 1993.
84. Giffin, A. and A. Caticha. *Updating probabilities with data and moments*. in *AIP Conference Proceedings*. 2007. AIP.
85. Guan, X., R. Jha, and Y. Liu, *Probabilistic fatigue damage prognosis using maximum entropy approach.* Journal of intelligent manufacturing, 2012. **23**(2): p. 163-171.
86. Skilling, J., *The axioms of maximum entropy.* Maximum-Entropy and Bayesian Methods in Science and Engineering, 1988. **1**: p. 173-187.
87. Haldar, A. and S. Mahadevan, *Probability, reliability, and statistical methods in engineering design*. 2000: John Wiley.
88. Mitchell, M., *An Introduction to Genetic Algorithms*. 1998: Bradford Books.
89. Kiefner, J.F. and W.A. Maxey. *The benefits and limitations of hydrostatic testing*. in *API's 51st Annual Pipeline Conference & Cybernetics Symposium, New Orleans, Louisiana (April 18-20, 2000)*. 2000.
90. Bao, Y., *Prediction of ductile crack formation in uncracked bodies*. 2003, Massachusetts Institute of Technology.
91. Bao, Y. and T. Wierzbicki, *A comparative study on various ductile crack formation criteria.* Journal of Engineering Materials and Technology, 2004. **126**(3): p. 314-324.
92. Bao, Y. and T. Wierzbicki, *On fracture locus in the equivalent strain and stress triaxiality space.* International Journal of Mechanical Sciences, 2004. **46**(1): p. 81-98.
93. Wierzbicki, T., Y. Bao, and Y. Bai. *A new experimental technique for constructing a fracture envelope of metals under multi-axial loading*. in *Proceedings of the 2005 SEM annual conference and exposition on experimental and applied mechanics*. 2005.
94. Wierzbicki, T., et al., *Calibration and evaluation of seven fracture models.* International Journal of Mechanical Sciences, 2005. **47**(4): p. 719-743.
95. Kausch, H.H., *Polymer fracture*. 2nd ed. Polymers, properties and applications. 1985, Berlin ; New York: Springer-Verlag.