Battelle
The Business of Innovation

Final Interim Report – Task 4.1

Compare - Contrast Analysis of Inspection Data and Failure Predictions versus Burst-Test Outcomes for ERW-Seam Defects

By

B. N. Leis
B N Leis, Consultant, Inc
517 Poe Avenue
Worthington, OH, 43085

Prepared for

U.S. Department of Transportation
Pipeline and Hazardous Materials
Safety Administration
1200 New Jersey Ave., SE
Washington DC 20590

Contract No. DTPH56-11-T-000003
Battelle Project No. G006084

June 23, 2013

Acknowledgments

Collaboration with Dr. J. B. Nestleroth now with Kiefner and Associates (then of Battelle) who working along with Ms. Smith and Mr. Robert Gertler of Battelle generated the burst-test data considered. Discussions with and comments by Dr John Kiefner also with Kiefner and Associates provided useful perspective on the capabilities of the more recent in-line and in the ditch inspection tools in contrast to his expectations. Useful discussions with Mr. E B Clark and the broader support of Ms. Jenny Smith, both of Battelle, also are gratefully acknowledged.

Executive Summary

This report is the deliverable for Task 4.1. The objective of this subtask was to quantify the effectiveness of 1) in-line inspection (ILI) and in-the-ditch (ITD) tools, and 2) predictive models used in integrity assessments in applications involving electric-resistance weld (ERW) seamed pipes with anomalies. Effectiveness was benchmarked against results from either full-scale burst-tests or field hydrotests.

Results reported as part of six project related subtasks are presented and evaluated in a compare-contrast framework that integrates case-specific data for three burst tests considering both inspection and failure pressure prediction. The six subtasks involved are:

- Subtask 2.1, which located and gathered ERW-seamed pipe anticipated to contain seam defects based on ILI and prior service history;
- Subtask 2.2, which inspected that pipe using ITD practices, and selectively involved ILI tool-pulls as the basis to prioritize pipe for burst-testing;
- Subtask 2.4, which assessed the utility of predictive failure models for defects;
- Subtask 1.3, which assessed the utility of ILI based on archival data;
- Subtask 2.3, which address small-scale testing to characterize ERW seam properties; and
- Subtask 2.6, which assessed approaches to characterize seam failures, including new/alternative technologies.

In addition to results from these subtasks, this report also integrates the results of extensive field hydrotesting and related ILI and ITD inspections of pipelines made using ERW line pipe. That work reflects the ILI of more than 1500 miles of pipeline since 2011, which has been broadly supported by field digs.

The viability and reliability of the tools was discussed relative to ERW seam features and the implications assessed in regard to the vintage as well as the modern pipeline systems. Thus, the results of this work help to define actions essential to improve the integrity management practices for ERW seamed pipe, with the possibility that those outcomes will have implications for standards development, or tool development and commercialization.

Results were presented for the use recent of second or third generation technology to detect and size axial seam defects, specifically spiral magnetic-flux leakage (SMFL) and electromagnetic acoustic transducers (EMATs). Trending for these results showed that much improved detection and sizing can be achieved today as compared to the outcomes developed early in the use of the first-generation inspection technologies. Trials using emerging ITD technology referred to as inverse wave field extrapolation (IWEX), which couples phased-array ultrasonic technology (PAUT) and time of flight diffraction (TOFD), were promising. With such technology indicated that step improvements are possible compared to currently available tools. That being said, given the limitations with the ITD technologies currently available the most reliable approach couples magnetic particle inspection with TOFD and PAUT.

The defects causing failure were found in the bondline, as well as in the upset region of the seams considered, which trace back to manufacturing setup and process upsets. Results developed showed that anomalies in the upset of the seam were much more stable than those in the bondline, which made clear that size alone does not define the threat posed by an anomaly. Metallography and fractography made clear the complexity of real seam defects, as compared to machined (idealized) features, not only in regard to their shapes and sizes but also in regard to microstructural differences that can affect failure response. It follows that there is a need to identify the location as well as the type of anomaly if such features are to be prioritized in condition assessment following the inspection.

It was found that reasonable predictions of failure pressure were possible for ERW seams when the shapes and sizes of the features were known, and the toughness local to the failure site could be estimated based on local properties data. This means that the models used to quantify failure pressure must be specific to the type of defect: that is bondline versus hook crack versus selective seam corrosion. While good predictions could be achieved when the differences in the severity of the features, and the local resistance to failure were addressed, scatter was evident when more rudimentary analyses were done based on nominal properties. As such, uncertainty in local toughness and ultimate tensile stress (UTS) can cause scatter in the predicted failure pressure, as can inadequate anomaly sizing.

Useful conclusions can be found throughout the report, of which the most important follow:

- ILI done using SMFL and EMAT tools focused in part on crack-like features associated with stress-corrosion cracking (SCC) over almost 1500 miles of liquid, highly volatile liquid, and natural gas pipelines made using low as well as high frequency ERW processes showed the technology to detect cracking has recently improved significantly. Based on data reported by the operator and their vendors –
  - Over the interval from 2008 to 2011, the probability of detection (POD) via EMATS for cracking due largely to SCC was found to be above 90% at a 95% confidence level, which is well above the normally cited POD of 80% at the same confidence level;
  - In contrast to failures on recently inspected lines using earlier generation technology, results specific to recent EMATs technology indicate that the probability of correct identification for lines with a statistically significant number of observations led to a success rate larger than 91% at 95% confidence level.
  - Likewise, in contrast to failures on recently inspected lines using earlier generation technology, such as transverse-flux magnetic flux leakage (MFL) tools, results specific to recent EMATs technology indicate that the success rate for probability of correct depth sizing has shown progressive improvements from 86% in 2008 up to 100% in 2011.

  - Because these results are in strong contrast to past experience and the expectations of some experts, there is a need to better understand and document the circumstances that underlie the improvements and more broadly replicate these observations.

- Collaboration between vendors and operators, and experts as needed, has contributed greatly to the improved detection and sizing capabilities;
- Vendor specifications for ILI tools were found in some cases to be equivalent to a 90% SMYS hydrotest, but this outcome was confined to specific combinations of line-pipe geometries, as for some geometries the tools were indicated to be less effective;
- The means to establish the viability of an ILI run via ITD technologies like phased array ultrasonic technology can be less reliable than desired;
- Limited testing with emerging ITD tools based on PAUT and related technologies indicated step improvements in anomaly sizing will be evident as compared to the status quo once such technology becomes commercially available;
- The irregular shape of real anomalies makes it difficult to quantify their size using the usual two parameters – maximum depth and length – which in some cases can complicate assessing the viability of ILI;
- Differences between the measurements from different sensor technologies are inevitable so long as complex features are characterized using in a few simple measurements – which also confounds assessing the viability of ILI;
- Failure is controlled by feature size and also the local properties, such that the interpretation of ILI and ITD tools must be taken in light of the features location, and the properties to the extent they can be inferred – likewise the development of inspection tools to routinely quantify local strength and toughness would affect a step improvement in failure pressure predictions;
- Meeting the challenge to "eliminate catastrophic failures in ERW pipe; as well as to the vintage system" is demanding, with continued improvement in both ILI and ITD technologies needed, including a focus on correctly calling the type of feature and its location – in addition to detecting and sizing it; and
- While the claim was made that inspection could be used to replace hydrotesting, and some success was noted that supports this view, it was also evident that in some cases the ILI equivalence that developed fell at or below code-minimum hydrotests levels – in addition it was evident that the effectiveness of detection as specified by the vendors is variable, depending on the pipe geometry and the properties within the ERW seam: key in this context is that the minimum specified detection in terms of anomaly size does not translate into constant ILI effectiveness when viewed in terms of equivalent hydrotest pressure.

Table of Contents

List of Figures

List of Tables

Page

List of Acronyms and Symbols

AIV	alternative integrity verification
AOS	A O Smith
CMFL	circumferential magnetic flux leakage (ILI tool)
DNV	Det Norske Veritas
D/S	downstream
EDM	electric-discharge machined
EMAT	electromagnetic acoustic transducer
ERW	electric resistance weld
FATT	fracture appearance transition temperature
FW	flash weld
HAZ	heat affected zone
HF (ERW)	high frequency (electric resistance weld)
HFI	high frequency induction
ID	inside diameter
ILI	in-line inspection
ITD	in-the-ditch (technology or method)
IWEX	inverse wave field extrapolation
KAI	Kiefner and Associates
LF (ERW)	low frequency (electric resistance weld)
MAOP	maximum allowable operating pressure
MDS	multiple data set(s)
MFL	magnetic flux leakage
MPI	magnetic particle inspection
NDE	non-destructive examination or evaluation
NSC	net-section collapse
NTSB	National Transportation Safety Board
OD	outside diameter
PAUT	phased-array ultrasonic testing or technology
PHMSA	Pipeline and Hazardous Material Safety Administration
PWHT	post-weld heat treatment
QA	quality assurance
QC	quality control
SMFL	spiral (helical) magnetic flux leakage (ILI tool)
SMYS	specified minimum yield stress
SMTS	specified minimum tensile stress
SSC	selective seam corrosion

TCPL	TransCanada Pipelines
TDW	T D Williamson, Inc
TOFD	time of flight diffraction
U/S	upstream
UT	ultrasonic testing
UTS	ultimate tensile stress
YS&T	Youngstown Sheet and Tube

Compare and Contrast Predictions of Inspections and Failure Models with the Outcomes of Burst-Tests of ERW-Seam Defects

Introduction

This report integrates the outcomes of prior work done in subtasks completed as required to meet the objectives of Subtask 4.1, which was to a compare predictions based on inspection data and failure models with reality as characterized by full-scale burst-tests for electric-resistance weld (ERW) seam defects. Specifically, this report builds on the outcomes reported for:

- Task 2.1[(1)*], which located and gathered ERW-seamed pipe anticipated to contain seam defects based on in-line inspection (ILI) and prior service history;
- Task 2.2[(1)], which inspected that pipe using in-the-ditch (ITD) practices, and made use of available ILI data supplemented by ILI tool-pulls to select and prioritize the available pipe for burst-testing;
- Task 2.4[(2)], which assessed the utility of predictive failure models for defects; and
- Task 1.3[(3)], which assessed the utility of ILI.

Also indirectly relevant to this subtask are the outcomes reported for Subtask 2.3[(4)], which address small-scale testing to characterize ERW seam properties, and Subtask 2.6[(5)], which assessed approaches to characterize seam failures, including new/alternative technologies. While the near-term results of this subtask help to define actions essential to improve the integrity management practices for ERW seamed pipe, those outcomes equally can have longer-term implications for standards development, although that is beyond the present scope.

Because prior reporting for Subtasks 1.4[(6)] and 4.2[(7)] has made clear that the types of defects that occur for high-frequency (HF) ERW seam processes (i.e., contact via shoes and noncontact via induction, termed herein high-frequency induction (HFI)) are also relevant to low-frequency (LF) ERW processes[1], the specific producer and vintage are not of concern. Accordingly, there is no distinction hereafter regarding LF versus HF processes, regardless of whether dealing with results from full-scale testing, or predictions of failure pressure, or ILI indications of defect sizing.

Consequently, this effort focuses on the viability of predictions based on inspection data and failure models in comparison with reality as characterized by full-scale burst-tests for ERW seam defects. Compare-contrast scenarios are presented hereafter for specific burst tests relative to predicted outcomes, with success or the reliability of predictions discussed along with potential causes for the issues evident in the laboratory-based inspection that appear to parallel those evident in prior reporting for the field data gathered during the course of this project[(e.g.,2,3)].

---

* Numbers in superscript parenthesis refer to the list of references collected at the end of this text.

[1] While not explicitly identified in this discussion, flash weld (FW) seams and those made via direct-current are LF ERW process – and considered within that grouping, although not explicitly called out in this discussion.

Significantly important in regard to the laboratory-based inspections is the limited sizes of the seam anomalies present in the pipe made available for evaluation. In general, such features were small (short and or shallow) and/or discontinuous, which effectively limits success in reference to both detection and sizing. Key in this context is the observation that such anomalies don't pose a practical threat unless they can fail under service-like pressures. Otherwise, limited success with such features suggests concern whereas in practice they would never pose a threat.

Finally, the first phase of this project was focused on archival data, which tends to reflect the use of what might be termed first or second generation technology[2]: these included an ultrasonic-technology (UT) based tool known commercially as Ultrascan CD, and a magnetic-flux leakage (MFL) based tool known commercially as the Transverse (Flux) Field Inspection (TFI) tool, which also has been referred to as circumferentially oriented MFL (CMFL). In contrast, later-generation tools are now available for use in a laboratory setting, which recently also have been deployed in the field. For this reason, the report closes with a review and analysis of results made available for recent commercial use of such technology. In particular, results are presented and discussed for spiral magnetic-flux leakage[(8)] (SMFL) and for electromagnetic acoustic transducers[(9)] (EMATs). As becomes evident, the advent of this next generation of technology opens to improved outcomes both in regard to anomaly detection and sizing.

Approach and Scope

As detailed in Reference 1, joints of pipe containing ERW seam anomalies have been sought and selectively gathered consistent with criteria chosen as part of Task 2 to support the objectives of this project. In total, seventy joints of pipe totaling about 2624 feet in length were collected, many of which had been subjected to ILI and ITD methods as part of the field rehabilitation work that removed them from service. Following receipt at Battelle, the pipe joints were inspected under laboratory conditions using ITD methods, with both time of flight diffraction (TOFD) and phased array ultrasonic testing (PAUT) used to further screen and size anomalies. Inverse wave field extrapolation (IWEX), which couples phased-array ultrasonic technology (PAUT) and time of flight diffraction (TOFD), also was briefly explored. In addition, joints whose diameter matched available SMFL tools also subject to ILI tool pulls as part of Task 2 of this project.

Where ILI or other inspection had been done, it reflects the use of the above-noted historically available tools for seam inspection. External nondestructive inspection methods also were typically used to evaluate the seam welds, both in the field prior to shipping, and as part of this project, to screen pipe for potential anomalies.

---

[2] First generation tools can be considered those developed in the 1960s and early 1970e, based on rudimentary MFL concepts, such as the early Linalog tools, while second generation tools can be considered the first-round of improvements on that technology, which appeared through the late 1970s into the early 1990s. What might be considered third generation tools thus reflect improvements in sensors and algorithms since that point in time.

Pipe joints were then prioritized for burst testing, with that prioritization based largely on the inspection results, and to a lesser extent also predictions of failure pressure. The failure pressure predictions played only a nominal role, as often the necessary pipe properties or details of the anomaly sizes remained ill-defined during the course of that prioritization. The burst testing was done to expose the defects causing failure, and to develop benchmarks against which to compare the effectiveness of the inspection and predicted failure behavior and defect sizing. The scope and the outcomes of such burst tests are constrained by the size, shape, and distribution of the anomalies actually present in the pipes supplied, as initially characterized in the field through an operator's rehabilitation project that necessitated removal of these and other joints of pipe. As becomes evident, in spite of first prioritizing pipe sought for the project based on field ILI and follow-up ITD inspection and then again using a range of ITD methods, these constraints are evident in the limited utility of some burst-test results. Details for each of the pipes evaluated including relevant mechanical and fracture properties, microstructures, chemistry and hardness are presented in Annex A.

The Burst Tests

Three burst tests were done on samples prioritized for such testing with the intent of exposing seam anomalies that had been indentified using ITD and in-line inspection, the details for which were covered in Reference 1. As noted in Reference 1, ILI results were available for two of three pipes considered. Because the planned purpose of this task was to compare and contrast the expectations versus actual results, this report adds to and selectively replicates details and images essential to document the outcomes of that process. Readers interested in other aspects of the testing and related activities should consult Reference 1.

Each burst test involved nominally steady pressurization at about 1 to 4 psi/second until failure occurred. In each test, the response at the anomaly anticipated to fail (based on absolute size determined primarily from the results of PAUT inspections done prior to the test) was captured on high-resolution digital video taken at 60 frames per second. Video capture was continuous, so there is no need to trigger a recorder: the memory simply dumps when full and capture restarts. Accordingly, the video provides a view of the failure within one-sixtieth (i.e., 0.0167) of second or less after failure initiates. It follows that which of the several features in a test pipe was the origin (i.e., actually triggers failure) can be directly determined or inferred simply by replaying the video.

Compare and Contrast – Failure Pressure Predictions

As indicated above, the results from such burst test provide quantitative evidence against which to compare the viability of failure pressure predictions, and the viability of the inspection method(s). This section considers the viability of failure pressure predictions, which for this reporting are based on the PRCI's pipe axial flaw failure criterion[10], PAFFC, which is one of several such tools[e.g., see 2]. The following sections consider the viability of the inspection tools to detect and size the features.

Test One: Pipe 16-16

The first pipe hydrostatically tested in this sequence of three burst tests was pipe 16-16, which was nominally 16 inches in diameter (actually 16.125 inches) with a wall thickness of 0.260 inches (actually 0.240 up to 0.252 inch local to the failure, due to over-scarf caused by misaligned skelp edges). The pipe was fabricated in 1960 of grade X52 steel.

According to Table 5 of Reference 1, 2 ILI calls, and 4 features noted therein as cold welds were detected in this pipe segment – all of which involved the ERW seam. The anomalies identified as ILI calls were detected by a CMFL tool in the ILI that motivated removal of this pipe segment. These anomalies were visually evident on the outside diameter (OD) of the pipe as minor surface offsets within or along the upset. These features were detected during the course of the ITD inspection done after receipt at Battelle, and noted in that context as possible hook cracks, as were smaller features noted as cold welds. Visual examination of the seam after it arrived at Battelle indicated that the ILI calls were mill defects apparently due to vertical skelp misalignment. The surface lengths of these anomalies as apparent on the OD were 6.25 and 7 inches, which respectively lay ~21 feet and ~23.5 feet from the U/S end of the pipe segment. Photographs of these features were included in Reference 1 as Figures 3c and 3d, respectively

Four additional anomalies were detected by in-the-ditch-methods, being identified as potential hook cracks. The two smaller anomalies were provided to DNV for use in their task directed at characterizing ERW seam toughness local to defect, while the two larger features remained in the segment that was pressure tested. One linear feature was detected at~6 feet from the U/S end of the pipe segment, and ~6 feet U/S from a girth weld. PAUT called this feature at ~3 inches long, with a maximum depth of 24% of the wall thickness. A second linear feature was found ~6 feet downstream (D/S) from that girth weld, or ~18 feet from the U/S end of the pipe segment. This feature was called at ~4 inches long with a maximum depth of 15% of the wall thickness. Photographs of these features were included in Reference 1 as Figures 3a and 3b, respectively

The burst pressure for this pipe segment was 1,690 psi[3], which corresponds to 98% of the specified minimum yield stress (SMYS) based on the local pipe wall thickness. This failure occurred without evidence of noticeable bulging, through the shorter of the two mill mismatch anomalies identified by the CMFL tool. The pipe split axially within the seam over a length of ~8 feet. As the split ran, it intersected the linear anomaly ~18 feet from the U/S end of the pipe segment and the mill mismatch that lay ~23.5 feet from the U/S end of the pipe segment. This split arrested well before the girth weld and as such the deeper of the features found via PAUT that lay ~6 feet U/S end of the pipe segment remained intact.

Initiation of the failure and the ensuing split was captured on video at 60 frames per second. The first frame captured from this video that showed evidence of failure, as shown in Figure 1, covers

[3] All failure pressures quoted are gage pressures, although not explicitly labeled as such.

almost the full length of the split. This figure indicates that a wispy mist formed as tiny droplets of water were released through the still tight axial split runs the full length of the split. On that basis, this ~8-foot long split developed within the interval between images, indicating that the speed of this fracture was at least 480 ft-sec^{-1}. This speed is within the range typically observed for dynamic fracture in high-pressure pipelines.

Figure 1 First frame of the video for Pipe 16-16

Because there was no evidence of bulging and the seam split formed within the time interval between images, which of the three features intersected by the split was the origin could not be simply identify from the video, or the presence of a local bulge. In more homogeneous materials it is reasonable to assume the speed of the split was uniform over its length, so the origin lies at its mid-length. However, because the ERW seam properties can be quite variable, this approach is not reliable.

Of the four possible origins, the deeper of the two crack-like features can be eliminated because it remained intact. The other three features all lay in close proximity to each other, within a length of about six feet. Visual examination of the fracture features did not indicate markers such as Chevrons, as occur with brittle fracture, so other features were considered as the basis to identify the origin. Steps along the fracture surface that show significant unidirectional undercutting can be useful in that context, as can other features that stand proud on primary fracture surface. Of these, undercutting was not evident, but there were features that indicated the mill mismatch feature located roughly mid-length along the split was the origin. In regard to Figure 1, early on in the failure process areas of locally dense (i.e., darker) mist indicate locally larger releases, and thus locally larger openings. As the size of the opening increases, the size of the droplets increases until water gushes out. On this basis, the area about mid-length in Figure 1 – highlighted by the dashed arc – was first to fail, which is consistent with the visual analysis of the fracture – which confirms the shorter of the two vertically misaligned features as the origin.

The crack-like anomaly that did not fail was removed and first assessed nondestructively using X-ray tomography[(e.g., see 5)], and then destructively using standard metallographic cross-sections made through the thickness in the vicinity of this feature. The intent of these studies was to better understand the observed failure response, and to quantify the OD-depth of the feature. Once the depth was measured, the section was etched with Nital to reveal the nature of the cracking relative to the microstructure and the bondline. Views of these cross-sections are shown in Figure 2.

The metallographic section indicates that in some places this anomaly involved cracking from both the outside as well as the inside diameter (ID) of the pipe, and that the cracking lay largely in or immediately adjacent to the bondline.[5] It should be noted that the fine cracking at the ID could have developed during the test, and so not have been apparent in the pre-test inspections. Both cross-sections show that the OD crack faces were separated slightly, which makes this feature a better reflector for the PAUT. The depth of the OD feature measured on the metallographic cross-section was 40% of the wall thickness, which is much larger than the maximum depth called by PAUT at 24%. This tomographic slice shows it to be slightly shallower than the actual measured depth, by which process it measures at ~36%. While such differences can embed some measurement uncertainty, the differences for the present appear to be driven more so by differences in actual crack depth. This is because the actual depth of the feature varies depending on the specific section where the measurement is made along its length.

a) metallographic section (Nital etch) **b) tomographic section**

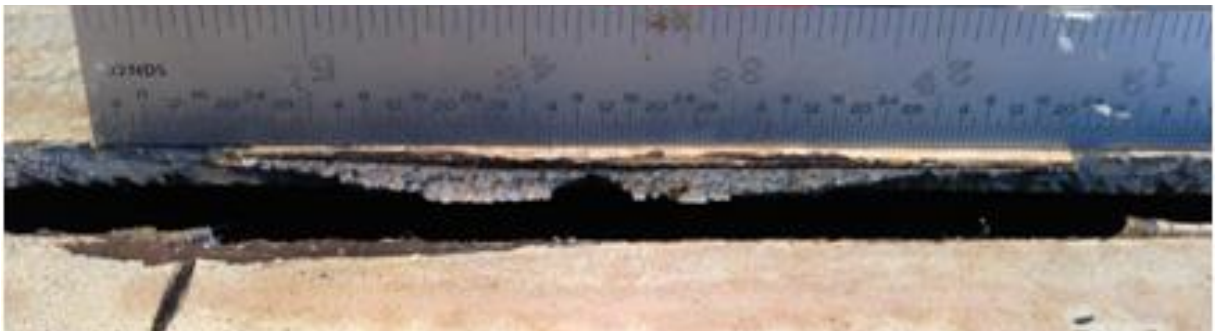
Figure 2 Transverse cross-sections that characterize the stable anomaly in Pipe 16-16

Such variations in depth were indicated by the full tomographic characterization[5], which showed the cracking to be discontinuous along its length. Discontinuous cracking could be identified by making serial metallographic sections through use of sequential grind – photograph cycles, but this is tedious. In contrast, discontinuous is more easily evaluated via fractography – if the cracking is deep enough to easily open the fracture in the laboratory, or if the fracture surfaces are exposed by the pressure test. Figure 3 illustrates the nature of the discontinuous cracking from the ID that was apparent over long segments of the split. The dark band along the upper edge of the portion of the split shown here is a shadow created by a shallow shear-lip, and so in this area is not indicative of a black oxide layer or other form or pre-existing anomaly.

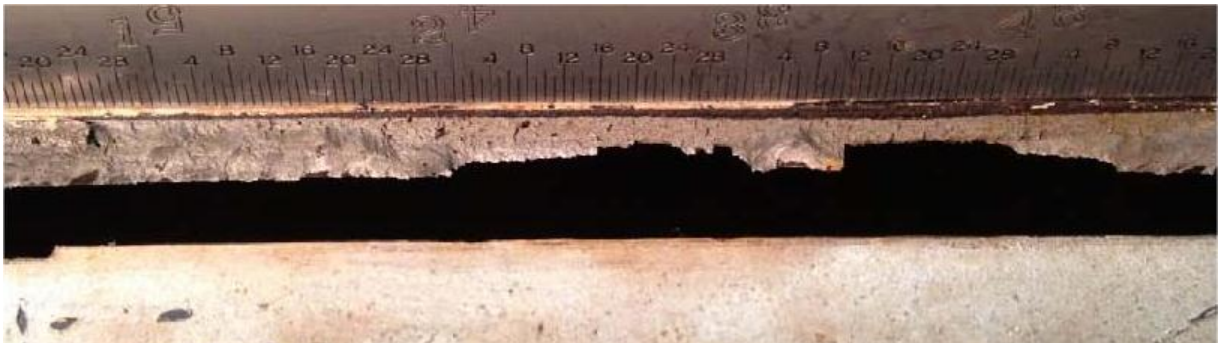
Figure 3 Discontinuous ID cracking along the stable anomaly in Pipe 16-16: t = 0.252 inch

It is apparent from Figure 3 that the discontinuous nature of this cracking shows axial spacing that is not too different from that for some stitched welds. But while similar in some ways to stitching, this oxide is present only part-through-wall. Each segment of this ID cracking contains a black high-temperature oxide, indicating this cracking formed as the seam was being made – as occurs for stitched welds. The remainder of the fracture surface shows areas of small reflective facets, which are indicative of localized zones of cleavage, and lower fracture resistance, which could give rise to a locally higher high fracture speed. Because this ID cracking is discontinuous, it is less likely to initiate cracking and serve as an origin as compared to a continuous anomaly.

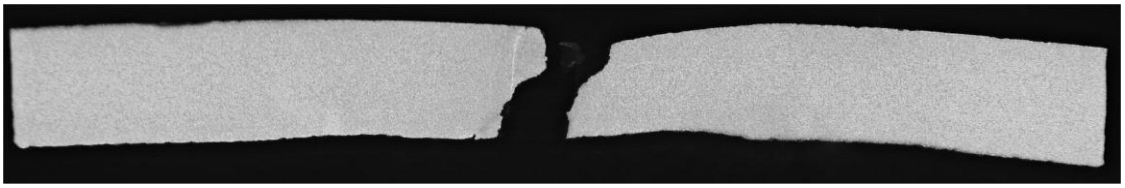
Figure 4 shows images of the likely fracture origin taken just after the burst test that have been taken from either side of the pipe to show each side of the fracture. In contrast to Figure 3, the darker band that runs along the upper edge of the view of the fracture surface shown in Figure 4a is not a shadow created by a shear-lip. Rather this is a discolored band that runs along the OD of the pipe, whose coloration is strong contrast to the fresh fracture below it, which occurred in the burst test.



a) view from the opposite side of to that shown in part (b



b) view from the opposite side of to that shown in part (a



c) cross-section through the seam within the origin (OD is up)

Figure 4 Views of the fracture surface at the failure origin for Pipe 16-16: t = 0.252”

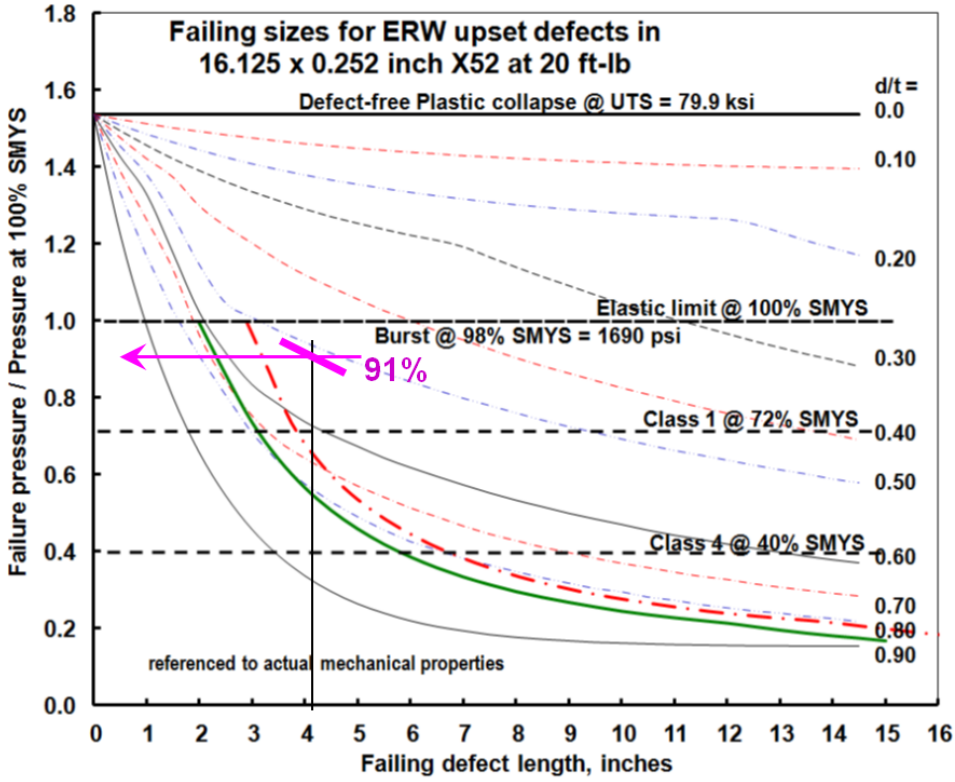
The continuous discolored feature evident in Figures 4a and 4b runs over a length of ~4.1 inches, and at its deepest is ~52% through the full thickness of the wall. The difference in the surface length of this misalignment measured on the OD at slightly more than six inches and that of the pre-existing defect noted above at 4.1 inches reflects the fact that the pre-existing defect lies well below the OD profile of the pipe.

In contrast to Figure 3, the black oxide layer on this pre-existing feature appears less intense. Figure 4c indicates this origin lies within the upset of the seam rather than in the bondline, although at this location as the cracking grew into the wall it shifts back toward and then runs parallel to the bondline. Etched higher magnification views indicate that the full-wall evident on the left side of the seam was formed by upset flow from the offset edge, which falls well below the OD profile of the pipe. Figure 4c indicates that the OD scarf trimmed the upset more or less even with the OD profile of the pipe, with a slight over scarf at the ID, which while not unusually large acts to further reduce the net wall thickness. The preexisting feature apparent in Figures 4a and 4b develops in this net wall, at the point where the wall is its thinnest. What is not clear without further study of the oxide layer is when this defect was formed.

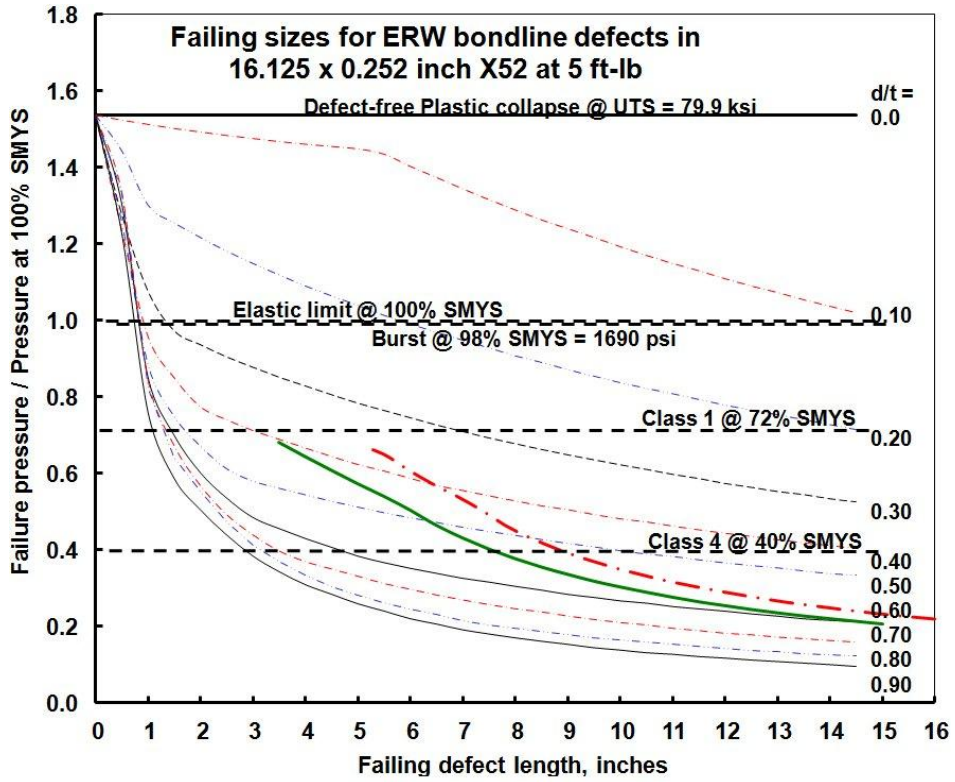
The dark thumbnail feature evident along the ID in Figure 4a is a shadow, which forms due to pullout of metal that contained a small surface anomaly that creates a step along the opposite face of the fracture, as is evident in Figure 4b. Figure 5 illustrates the area of the pullout in more detail, with the overall shape of the thumbnail-shaped step on the fracture surface highlighted. Within the highlighted field lies a small thumbnail that is covered by black oxide, which also has been highlighted to help identify it. This thumbnail shaped area of black oxide and the surrounding material on this pullout-step lies in the bondline for this ERW seam. The presence of this small pre-existing defect on the ID evidently is responsible for the step in the ID crack path from the seam into the bondline – but as this is a small step that is due to a trivial bondline defect, this apparent origin for the failure forms due to the much larger OD mill mismatch and related pre-existing defect, which fall within the upset.

Figure 5 Detail in the area of the pullout evident as the shadow-thumbnail in Figure 4a

Predictions of the failing defect sizes for Pipe 16-16 are presented in Figure 6 based on Release 5 of PAFFC (Pipeline Axial Flaw Failure Criterion)[10], which was developed at Battelle under funding from the PRCI. These predicted failure boundaries are based on the actual pipe size and its properties, which are presented in Annex B. Note that Figure 6a presents results at 20 ft-lbs as representative of failure in the upset of the seam, while Figure 6b presents results at 5 ft-lbs as representative of failure in the bondline. These full-size equivalent (FSE) Charpy v-notch



a) seam features in the upset



b) seam features in the bondline

Figure 6 Predicted failure boundaries for Pipe 16-16

(CVN) energy (toughness) levels correspond to energies that have been developed specific to the upset and bondline at a temperature relevant for the burst test, which have been linearly scaled from sub-size results based on energy density.

The y-axis in Figure 6 is the predicted failure pressure normalized relative to the pressure that corresponds to SMYS. Thus, a value of one on the y-axis corresponds to the pressure at 100% of SMYS, or nominal yielding. The x-axis in this figure is the axial length of the anomaly while the contours shown represent anomaly depth normalized by the wall thickness. Depth contours are shown at intervals corresponding to one-tenth of the wall thickness, such that anomaly depth must be interpolated relative to any pair of contours, as must the length. As just noted, values of constant failure pressure appear as horizontal lines in such a plot. The uppermost horizontal line corresponds to plastic collapse in a defect free pipe, and so by definition lies at the ratio of the ultimate tensile stress (UTS) to the yield stress.

The failure pressure as predicted in Figure 6 reflects the lowest pressure that is independently calculated for each of the two known failure processes in line pipe – plastic collapse and fracture. Failure due to plastic collapse is controlled by the UTS, while that for fracture is controlled by toughness, where the resistance relevant to the failure site can be modified by local constraint effects. In the format of Figure 6, depth contours that depict failure by plastic collapse through the net ligament (net-section) below a defect (i.e., net-section collapse, NSC)) are evident as 'nested' contours. The spacing between these contours is proportional to the difference in wall thickness, such that for very long defects the failure boundaries are spaced uniformly in proportion to the depth, d, divided by the wall thickness, t. Sharp changes in the shape of the failure boundaries are evident for some of the failure boundaries shown in Figure 6 along the otherwise smooth nested trends. These sharp changes in shape correspond to the transition from collapse to fracture controlled failure. Because fracture can occur before the pipe wall develops NSC, such breaks along the smooth nested boundaries lead to lower failure pressures as compared to the NSC trend.

In reference to the fracture resistance for Pipe 16-16 as characterized by the FSE CVN energy at the temperature for this burst test as quantified in Annex A, the bondline toughness can be taken at ~5 ft-lb while that for the upset is about 20 ft-lb. On this basis, Figure 6a can be used to estimate the failure pressure of the OD mill mismatch (Figure 4) that grows into the upset, while Figure 6b can be used to estimate the failure pressure of the small ID cold weld feature (Figure %) that grows in the bondline.

It is apparent from Figure 6a that at a CVN energy of 20 ft-lb only very shallow and very deep features achieve NSC, as the failure boundaries at d/t = 0.1 and 0.9 match the smooth nested format of NSC. On that same basis, Figure 6 indicates that all shallow features with $d/t \leq 0.3$ and lengths less than about 7 inches are indicated to fail by NSC at any seam toughness level at or above 20 ft-lb. Parallel analysis indicates that for any seam toughness level above 10 ft-lb, all shallow features with $d/t \leq 0.2$ and lengths less than about five inches are indicated to fail by NSC. If the toughness is further reduced to 5 ft-lb or above, which represents this bondline, then

the sizes failing by NSC are significantly reduced, as then only shallow features with $d/t \leq 0.1$ and lengths less than about five inches are indicated to fail by NSC. Continued reduction in toughness affects a reduction in the sizes that fail by NSC, such that at 3 ft-lb and below all features with $d/t \geq 0.1$ fail by fracture, while shallower features could still fail via NSC.

In addition to the failure boundaries just discussed, Figure 6 also includes trends that indicate the onset of through-wall collapse, and its instability, which are shown respectively as the trends that run from the values of the x-axis near one diagonally toward the lower right corner of the image. Combinations of pressure and defect size that lie above the dashed boundary are indicated to rupture, while those that lie between them are transitional cases tending to instability. Combinations that lie below the lower of these trends are indicated to remain stable. If PAFFC correctly captures the failure process, then the failure of the small black oxide covered feature just visible in Figure 5, which lies in the bondline, develops by NSC at a pressure approaching the defect free pipe response. In contrast, the much larger discolored feature associated with the mill-mismatch that runs within the upset of the seam over a length of ~4 inches with a maximum depth of ~52% through-wall (Figure 4) is predicted to fail about 91% of SMYS and rupture. The actual pressure relative to the local wall was 98% and resulted in rupture. Given the variability that can develop along an ERW seam, the predicted failure reasonably matches the actual response, being conservative relative to the actual failure response by about 7%.

The operator that donated Pipe 16-16 has done extensive burst testing as they completed their rehabilitation of the subject pipeline. This data was provided in support of this project[11], and available for consideration in the work since the project began. This reporting included a field estimate of defect size based on the features exposed in the testing, and photographs generally useful in assessing whether failure initiated in the bondline or the upset. Considering that the data for Pipe 16-16 in Annex A can be considered typical of the other pipe joints tested from the same pipeline, then based on that data the utility of current predictive tools can be more broadly assessed.

Figure 7 presents the results of such predictions. The x-axis in this figure is the actual failure pressure normalized by SMYS, while the y-axis is the predicted failure pressure normalized in the same way. It is apparent from these results that the predictions tend to be conservative, and that there is significant scatter. Two factors drive scatter in predictions – uncertainty in the properties and uncertainty in the defect size and its position within the seam. In reference to Figure 6 it is clear that uncertainty in each of these aspects can cause very large swings in failure pressure, particularly when dealing with defects less than a few inches in length, which becomes extreme when they are in the bondline. For example, the same 3-inch-long 30% deep feature in the upset fails at 134% of SMYS while in the bondline it fails at just 88% of SMYS, for an error of over 50% relative to failure in the bondline. Variations in toughness or the UTS can have a similar impact for failure controlled, respectively by fracture versus NSC. Small errors in length can affect significant errors – particularly for shorter features a few inches long, as the failure boundaries vary strongly with defect length for such cases. But if a feature that that comprises

two or more thumbnail segments is miscalled or misrepresented as a single feature, the error in failure pressure can become very large. Overlaid on this variability is the uncertainty in toughness that can trace to variations in the burst test (or hydrotests) temperature. In regard to the toughness trends shown in Figure A3 in Annex A, it is apparent that the ductile-to-brittle FATT for the seam tends to lie within the range of typical field-supplied water. An uncertain or varying burst/hydro test temperature thus can cause significant variation in predicted failure pressure when as in the above predictions a constant temperature is adopted for all predictions. It follows that disparities in predicted versus actual pressure the order of 10% are not surprising.

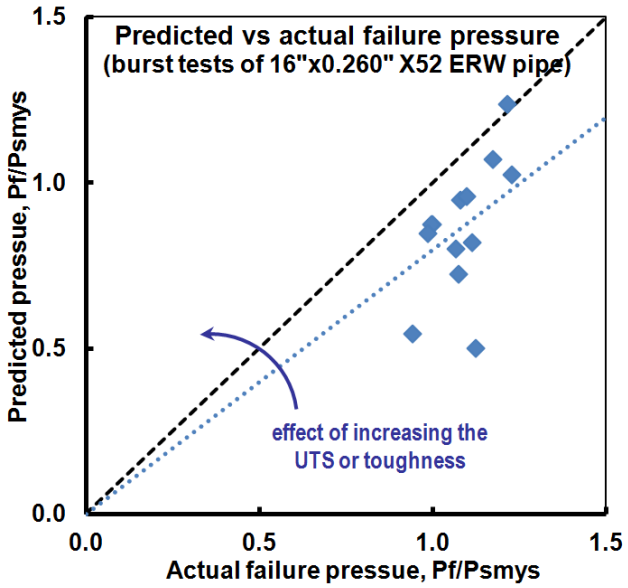


Figure 7 Failure pressure predictions for field testing of 16” x 0.260” X52 ERW pipe

In regard to Figure 7, the results would be improved through consideration of a slightly higher resistance in contrast to that for Pipe 16-16. Variability in the predictions enters as well in that the temperature in for the field testing was not always recorded. The sensitivity of the predictions to toughness noted above in regard to Figure 6 underscores the need to better quantify this parameter in order to more effectively validate predictive models. While such uncertainty could be addressed by embedding conservatism in predicted failure pressure, the effect of that conservatism causes non-conservative predictions of the related defect size, which confounds analysis of re-inspection intervals. It follows that models that predict failure pressure as a function of anomaly size must be accurate and embed minimal conservatism. And given the difference between bondline versus upset properties, it also follows that inspection tools must not only better size seam anomalies, but also must better locate them relative to the bondline versus the upset.

Test Two: Pipe 10-07

The second pipe subjected to burst testing was pipe 10-07. This pipe was nominally 10 inches in diameter (10.313 inch actual diameter) with a nominally 0.231 inch-thick wall thickness (0.235 inch actual thickness local to the failure) made of grade X52 steel. The burst pressure

corresponding to SMYS for this grade and pipe size is 2288 psi. There was no field data available for this pipe developed by either ILI or ITD methods.

The laboratory ITD technology detected two indications that were noted as likely in bondline, as has been detailed in Reference 1. The anomalies were sized at 3.3 inches long with a maximum depth equal to 29.8% of the wall thickness, while the second was 1.2 inches long with a maximum depth equal to 19.1% of the wall thickness.

The burst occurred at 3050 psi, which for the steel grade and pipe geometry involved corresponds to 133% of SMYS. In spite of the relatively large sizes of the two anomalies detected using the ITD tools, the pipe failed in the seam in an area remote to where anomalies were detected, ~5 feet from the downstream (D/S) girth weld. Unfortunately, because the pipe burst remote to either of the detected features, which were the focus of the video, the burst was not captured on the video. However, the only candidate origin was readily identified following the test as an ID defect.

The resulting split ran over a total length of about 6 feet, suggesting that as for the prior test the seam had relatively low toughness over the distance that it propagated. Because that the split ran just 3 feet either direction, the other anomalies detected using the ITD methods were not exposed. Given that the fracture surface over this full length appears more or less similar, it is likely that the arrest of the split reflects decompression of the pipe, which happens rapidly when water is used as the pressurizing media.

Figure 8 shows a portion of the fracture surface that contains the only candidate origin. As shown there, this origin involved a few patches of black oxide connected to the ID surface. In addition, there are a few shadowed areas that are not oxide covered. Based on this and other observations the failure originated in the bondline. These oxide covered patches (cold welds) are contained within a largely featureless fracture surface that runs to about half-through-wall at its maximum depth, and over a length of ~2.5 inches. The largest of the patches contained within this featureless fracture surface has a surface-breaking length of ~0.355-inch, and a maximum depth of ~0.094-inch, or 42% of the wall thickness. Examination of the surrounding fracture features indicates that this black oxide patch was the triggering defect – which is logical given that it is clearly the largest of these oxide patches.

a) overview of the origin **b) oxide covered defect**

Figure 8 Cold-weld features at the origin for the failure in Pipe 10-7: t = 0.231”

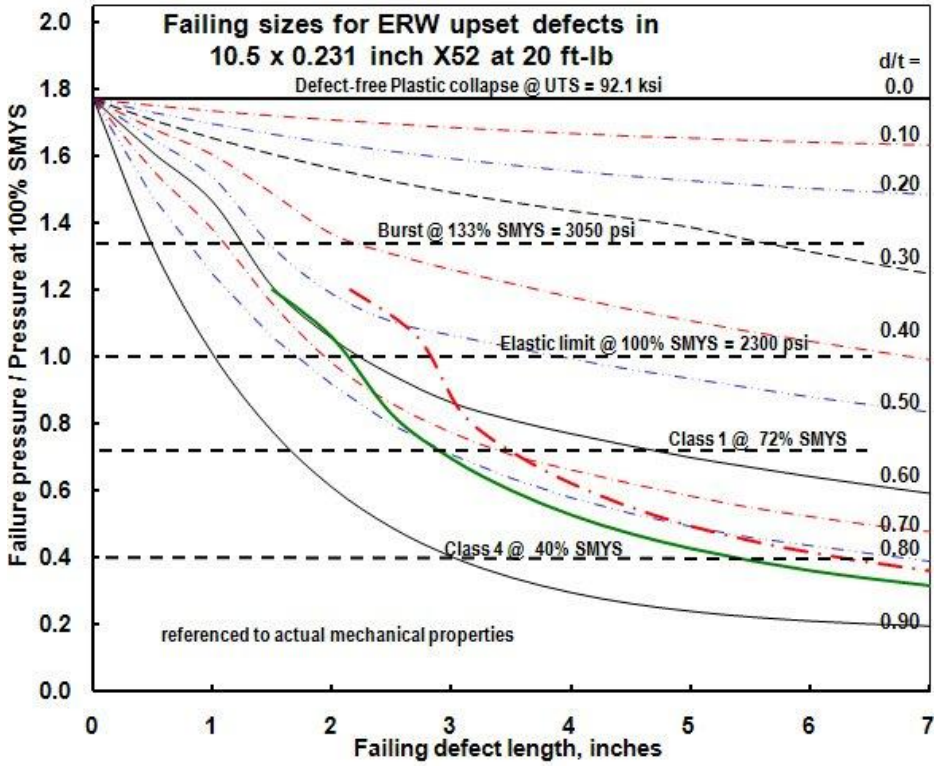
In reference to the fracture resistance for Pipe 10-7 as characterized by the FSE CVN energy at the temperature for this burst test as quantified in Annex A, the bondline toughness can be taken at ~3 ft-lb while that for the upset is again about 20 ft-lb. On this basis, Figure 9a can be used to estimate the failure pressure of the anomalies in the upset the worst of which was 3.3 inches long with a maximum depth equal to ~30% of the wall thickness. Figure 9b is relevant to the small cold weld feature (Figure 8) that grows in the bondline.

The format of Figure 9 is as it was for Figure 6. Thus, the predicted failure pressure reflects the lowest pressure that is independently calculated for each of the two known failure processes in line pipe – plastic collapse and fracture. All of the same general statements made in regard to the trends shown for Pipe 16-16 apply in regard to Figure 9. Failure due to plastic collapse can be identified relative to 'nested' contours that depict NSC through the ligament below the defect. Where sharp changes are evident in the otherwise smooth nested boundaries correspond to the transition from collapse to fracture controlled failure. As noted above, because fracture can occur before the pipe wall develops NSC, such breaks along the smooth nested boundaries lead to lower failure pressures as compared to the NSC trend. As for the results for Pipe 16-16, the results in Figure 9 show a very strong effect of toughness on failure pressure and the potential for rupture in lieu of a leak. Comparing part a) that represents the response of anomalies in the upset versus part b) that represents the response in the bondline, quite large anomalies can develop in the upset and remain stable relative to the response of much smaller anomalies in the bondline that fail, and can lead to rupture because of their very high failure pressures.

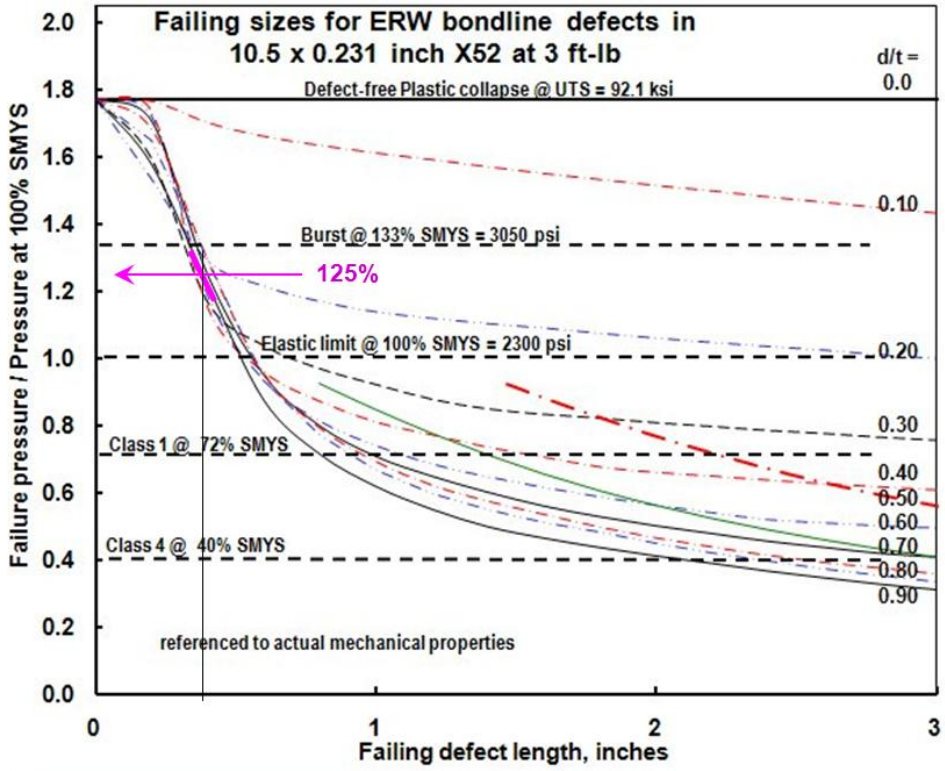
Inspection of Figure 9b indicates that an anomaly with a length of 0.355-inch and a maximum depth of 42% of the wall thickness fails at a pressure of ~125% SMYS, and is indicated to rupture – although the trends shown for the bondline are limited to nominally elastic response. This predicted pressure is somewhat less than the actual value, which was 133% SMYS. As above, this prediction is slightly conservative, erring by about 6%. At this pressure, features in the upset remain stable at lengths less than ~4 inches at depths ≤ 30% of the wall thickness. In particular, Figure9a indicates that the largest feature in the upset fails at a pressure approaching 150% SMYS, which is well beyond the actual failure pressure. Thus, it is not surprising that the relatively longer and quite deep features located by the ITD technologies in the upset remained stable, because the local toughness was much higher there than in the bondline. It follows as expected that failure is controlled by feature size and also the local properties. Thus, the interpretation of ILI and ITD tools must be evaluated in light of the feature's location, and the properties to the extent they can be inferred.

Test Three: Pipe 22-11

The third pipe hydrostatically burst was Pipe 22-11. This pipe was nominally 22 inches in diameter with a 0.281 inch-thick wall (actually 0.290 inch local to the failure) that like the other two pipes was made of grade X52. The burst pressure for this grade and pipe size corresponding to failure at a wall hoop stress equal to SMYS is 1371 psi. The actual burst pressure for this pipe was 1850 psi, or 135% of SMYS. The fracture ran approximately 6 feet. As was the case for



a) seam features in the upset



b) seam features in the bondline

Figure 9 Predicted failure boundaries for Pipe 10-7

Pipe 10-7, failure initiated at a site remote to either of the detected anomalies, which for this case occurred 5 feet from the D/S girth weld.

As for the prior test, the burst was not captured on video because the failure initiated away from the known anomalies, as detailed in Reference 1. EMAT ILI data for this pipe provided by the operator indicated two 3-foot-long seam anomalies, and many smaller seam anomalies. The use of the ITD tools at Battelle confirmed these results, which the ITD tools identified as potential bondline features. Post-test examination of the pipe confirmed the extent of the crack-like indication, with metallographic sectioning used to locate its position relative to the seam, and measure its depth.

Figure 10 captures the details evident in a metallographic section, for which the bondline is just visible at the left margin of the image. This figure confirms the presence of a surface-breaking feature whose appearance at the OD surface is crack-like. While the feature appears crack-like from the OD of the pipe, the cross-section indicates that it lies at about 45° to the pipe's surface, and that terminates in a very blunt tip. High-power optical microscopy did not show evidence of microcracking within this feature, and while not oxide-filled it does show their presence. The shape of this feature is not inconsistent with what Battelle's archives have termed arc burns, while in other cases such features also have been referred to as folds or laps (apparently due to skelp production). While past cases of similar features termed arc burns were present on sides of the seam, this feature occurred only on one side.

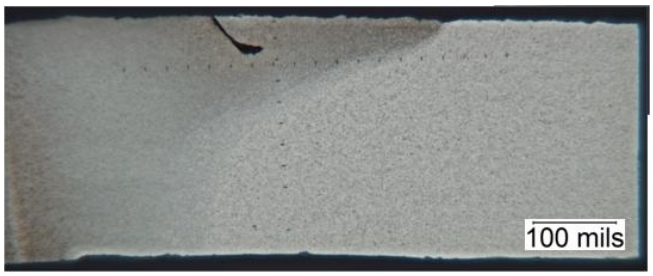


Figure 10 Metallographic section through the seam in Pipe 22-11: t = 0.290"

Hardness traverses were run, as evident in their tracks in Figure 10, to better assess the significance of the microstructural variations evident of in this figure. Figure 10 indicates that the heat-affected zone (HAZ) of the seam is modified and shows traits that could be due to seam normalization. This is consistent with both the vintage and supplier of the pipe involved, which involves mid-50s construction and production by Youngstown Sheet and Tube (YS&T).

The hardness traverses indicated that the Knoop (500g) hardness levels were all less than 234 on the vertical path, with nothing over 248 being found on horizontal path. These levels correspond on the Vickers scale to 223 and 237, while on the Rockwell scale they translate to R_B at 99 and R_C at 20 – so there is nothing "hard" about this microstructure. This is consistent with the absence of microcracks and a seam that appears to have undergone PWHT. Thus, while somewhat different than the base metal, it is not radically different. Such hardness levels suggest that the UTS will be well above that of the base metal, while due to the somewhat coarser

microstructure the FATT will be reduced somewhat from that of the base metal – although this is only a consideration if failure occurs outside the seam and fracture controls that failure.

Based on Figure 10, this long possibly discontinuous axial feature lies well outside the upset zone in the base metal, and has a maximum depth into the thickness of about 16% of the local wall. If continuous, its total length was ~83 inches.

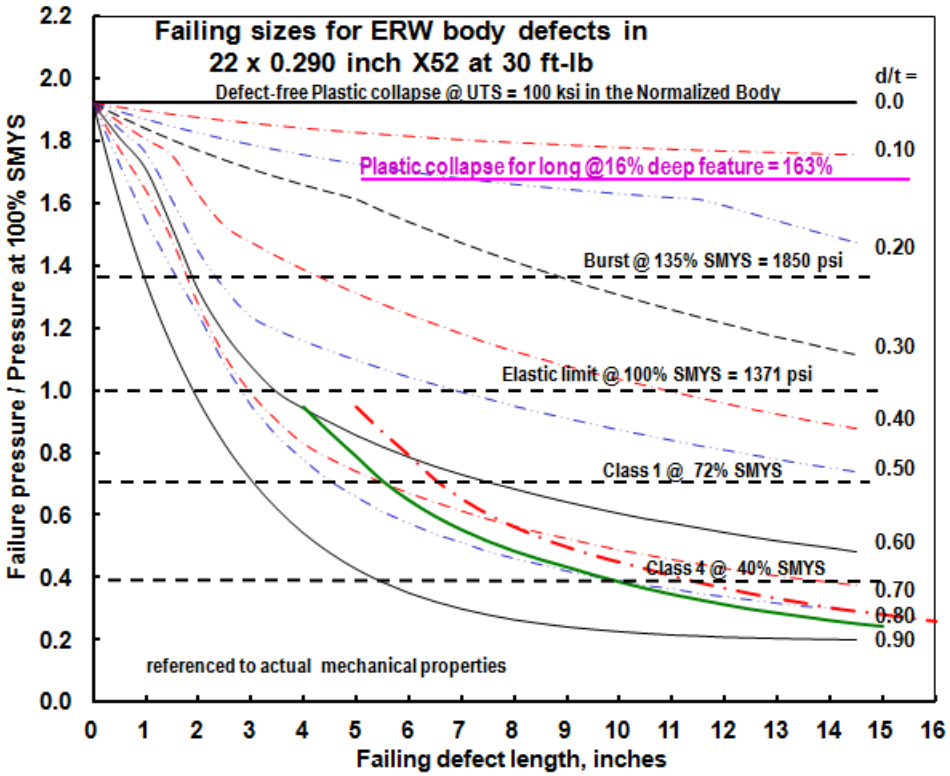
In contrast to the expected control of failure by the 83-inch long feature, failure occurred at the short feature shown in Figure 11. Examination of the features exposed along the seam indicated that the initiating anomaly lay at the ID. If the component parts of this feature are considered to interact, then the overall length of this feature is 0.384 inches. However, it the component parts of this feature do not interact, then the length is that of the longest of these parts, which indicates the length is 0.234 inches. Regardless of the length adopted, the maximum depth is 0.097-inch, or 33% of the wall thickness. This short feature lay within a smooth fracture surface that ran from the ID of the pipe to about mid-wall over a length of about 2.5 inches. The long anomaly that was detected by the ILI and ITD tools but did not fail was removed and assessed destructively, the results of which have been discussed above in regard to Figure 10.

a) overview of the origin (ID surface is down) **b) oxide covered defect**

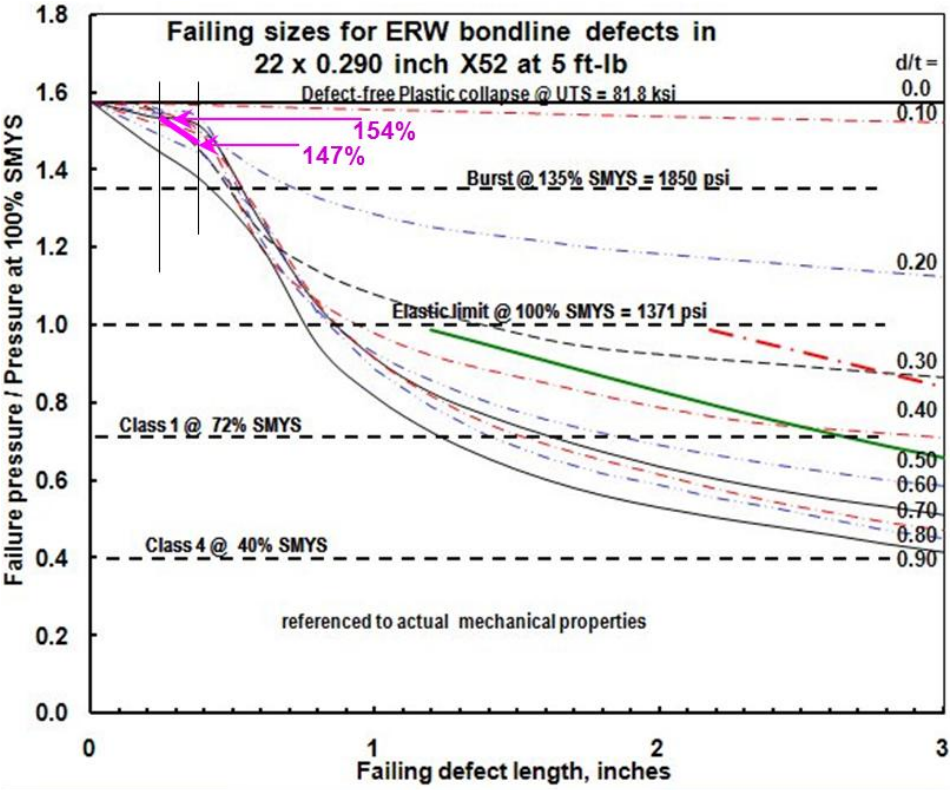
Figure 11 Views of the cold welds at the origin for the failure in Pipe 22-11: t = 0.290”

As for the other test pipes, predictions have been made of the failure pressure based on the results of PAFFC, which are presented in Figure 12 based on the toughness data presented in Annex A. As above the format of Figure 6 that presents the combined results for both plastic collapse and fracture is adopted. Thus, the same general statements made in regard to the trends shown initially for Pipe 16-16 and considered for Pipe 10-7 in regard to Figure 9 apply equally here.

It is apparent from Figure 12 that as for the prior discussion the results for Pipe 22-11show a very strong effect of toughness on failure pressure and the potential for rupture in lieu of a leak. In contrast to the prior discussion that contrasted the response of defects in the bondline versus those in the upset, the possibly continuous 83-inch-long stable feature was not in the upset. Rather, it lay in the pipe body, within a microstructure that was marginally harder than the base metal. As such, the present discussion contrasts the response of defects pipe body with those in the bondline. In that context, Figure 12a presents the response of anomalies in the pipe body in comparison to those in the bondline that are shown in Figure 12b. Note that the UTS for the prediction of the pipe body feature reflect its presence in a locally harder microstructure, with correspondingly higher UTS. Thus the y-axis scales differ between Figures 12a and 12b.



a) seam features in the pipe body



b) seam features in the bondline

Figure 12 Predicted failure boundaries for Pipe 22-11: t = 0.290"

It is apparent in comparing Figures 12a and 12b that quite large anomalies can exist in the pipe body and remain stable relative to the response of very much smaller anomalies in the bondline, which can fail and possibly lead to rupture because of their high failure pressures.

Inspection of Figure 12b indicates that an anomaly with a length of 0.384-inch and a maximum depth of 33% of the wall thickness fails at a pressure of ~147% SMYS, and is inferred to rupture. This same process indicates that the slightly shorter component of this feature whose length is 0.234 inch would fail at about 154% SMYS given the same maximum depth and also is inferred to rupture. Both predictions are very close to the actual failure pressure that was 135% SMYS, but are nonconservative. Because the possibly continuous 83-inch long feature lies in a microstructure harder than the base metal for this pipe, the failure response for the base metal is underestimated by the properties reported in regard to Figure 12b. As noted above, this difference has been accounted for in the properties used in regard to Figure12a, which uses a lower bound estimate of local UTS at 100 ksi. On this basis, the response of this long feature was determined from NSC analysis and the value of the UTS inferred consistent with the local hardness. The NSC failure stress can be determined for long defects simply by the ratio of the net wall thickness to the full wall multiplied by the local UTS. Such a calculation indicates that this long defect remains stable at the test pressure, and survived because the local strength was much higher than for the bondline, and toughness did not control failure.

Thus, as for Pipe 10-7 and true in general, failure is controlled by feature size as well as the local properties, which in turn suggests that the interpretation of ILI and ITD tools must be taken in light of the features location, and the properties to the extent they can be inferred.

Summary for Pressure Predictions

Table 1 summarizes the failure pressure predictions for the three burst tests.

Table 1 Summary of failure pressure predictions

Pipe	Failure pressure, % SMYS		Failure details	
	Actual	Predicted	Location	Type
16-16	98	91	upset	hook-like
10-7	133	125	bondline	planar
22-11	135	147	bondline	planar

Note from this table that failure for two of the three burst tests completed involved pressures in excess of 130% of SMYS. For X52, the specified minimum tensile stress (SMTS)[(e.g.,12)], which defines the minimum value for the UTS reported in Annex A, is 66 ksi, such that the ratio of these specified minimum properties, SMTS/SMYS, is 66/52 = 1.27. Thus, two of the three burst tests achieved pressures in excess of that associated with SMTS for the pipe. The third pressure test achieved a local pressure corresponding to 98% SMYS, which is well above the code[(e.g.,13)] minimum maximum hydrotests pressure to establish fitness for service of a pipeline. It follows that each of these tests failed at pressures well above levels considered practical.

While the test results are limited, when taken with the outcomes of Reference 2, which more broadly reviewed the available models to predict failure pressure, and the results shown in Figure 9, the outcomes in Table 1 indicate that the failure pressure can be estimated if the properties can be reasonably estimated, and the defect locations and sizes are known. However, significant scatter exists in regard to the results presented, which indicates that available models have not been adapted adequately to the unique aspects of ERW seam failures. Plans are in place as a part of Phase II of this project to begin to address the differences and adapt the models, but it is conceivable that work could be needed beyond that first attempt to address these aspects. Of importance to the discussion that follows is the fact that defects in the pipe body respond differently than those in the upset, or those in the bondline. It follows that the need to identify the feature type and its position relative to the bondline is an essential output from any inspection – which to this point remains a challenge.

Compare and Contrast – Inspection Aspects

As the content of Reference 1 details, a total of three burst tests were done to expose the seam anomalies that were identified via the ITD inspection and also ILI for two of the test pipes, which in this section are considered in light of the viability of these inspection methods. The results presented herein indicate that when the sizes of the defects are known, along with viable estimates of the properties failure pressure can be predicted. As such, failure response can be reasonably characterized as a function of pipeline condition and line pipe properties. This establishes the viability of one fundamental tenant of PHMSA's approach for pipeline integrity management. However, it was clear that such predictions show strong variability that is tied directly to uncertainty in both the sizing and the properties. In turn, this makes the successful deployment of this approach to integrity management dependent on the ability to consistently measure properties, and to detect and size anomalies, and place them in the bondline, the upset, or the pipe body.

It is apparent from the prior section that the inspection technologies failed for two of three cases considered. While this presents some concern regarding PHMSA's approach for integrity management, it must be recognized that the seam anomalies initially available for this evaluation were much smaller than those that pose an integrity threat under practical circumstances. Because all such tools have detection thresholds and their relative sizing capabilities improve as the size increases, the small size of the anomalies available conspired against success. As discussed earlier, Table 1 shows that failure occurred in the burst tests at pressures well above that practical, even in so called high-pressure spike testing for two of the three cases considered, while the third failed well above the minimum Code-required hydrotests pressure. Thus, the results developed by this testing reflect pressures well above any practical operating circumstance, including that for grandfathered pipelines. For this reason, the issues evident with the ILI and ITD tools must be reconsidered in light of trends developed for anomalies whose size poses an integrity threat under practical circumstances. This aspect is considered next, and will be more broadly addressed in Phase II.

In contrast to the smaller features exposed in the burst tests done as a part of Task 2.2, the work reported under Task 1.3[3], which assessed the viability of vintage ILI and ITD tools, assembled a broad database for those technologies applied to in-service pipelines. Because Task 1.3 dealt with in-service pipelines that had experienced incidents and with anomalies that were exposed in hydrostatic retesting, the sizes of the anomalies are relevant to pipeline condition assessment via ILI or ITD technologies. Results reported for Task 1.3[3] cover 13 ILI crack-tool runs through pipeline segments constructed with ERW pipe, which are compared to results developed from excavations and direct study of the anomalies located by the tools.

The archival data evaluated in Reference 3 represented the use of three different types of ILI technology, which given the timeline addressed by this work reflect first-generation crack-detection tools. Nine of the 13 inspection runs involved UT angle-beam detection the results for which represent two vendors, three involved CMFL runs by two vendors, while the last involved one EMAT tool. As noted in Reference 3, these inspections covered 741 miles of liquid, highly volatile liquid, and natural gas pipelines made using LF ERW and/or HF ERW processes.

In the words of Reference 3, "these cases illustrate that ERW seam integrity assessment via ILI is still a work in progress," which reflects the observation that disparities were evident between the ILI calls and the features evident by way of other detection and sizing practices. Where such observations developed in regard to the use of field tools for non-destructive examination (NDE) to size the features, the results might be rationalized in regard to issues in the use of the specific ITD technologies used. However, in many cases the benchmark for comparison came via direct methods for detection and sizing, by a subsequent hydrostatic test or burst testing of excised pipe, and follow-on characterization of the anomalies. It was stated that these direct comparisons show that while the tools often found anomalies, other anomalies that where supposedly within the detection capabilities of the technologies were mischaracterized or entirely missed. It was also evident that unless carefully calibrated the use of field NDE to size the anomalies located by the tools not reliable.

The results of Task 1.3 reflect the same type of commercially available first generation tools and practices that underlie the burst test results developed in Task 2.2. It is apparent in this context that the limited size of the features evaluated in Task 2.2 was not directly responsible for the inconsistent sizing. Recognizing that the work done under both of these subtasks reflects first generation technology for crack detection and sizing that dates back to the 1990s, it is instructive to explore the extent to which improvements have occurred since the introduction of those early tools well over a decade ago. Accordingly, results were sought to quantify some of the more recent developments. Specifically, information was sought for what is the first-generation release of SMFL technology, and what might be considered nearly third-generation EMATs technology.

Results for the SMFL tool reflect early commercial runs of this technology, which have been kindly provided by T. D. Williamson[14] (TDW). A few years back, TDW noted[15] the promise of the technology in seam-weld applications, referring to its "fusion with traditional high-

resolution MFL for improved classification of seam-weld anomalies." It potential utility in that context makes this first-gen development highly relevant to this reporting – in spite of its potential for a second or third development cycle. More critically, because it is a "combination tool" that couples many technologies, it can provide utility beyond seam inspection for cracking. In contrast, the EMATs concept has been explored for crack detection for about two decades, and as noted above is nearly a third-generation technology. Results presented relative to the EMATs tool reflect commercial inspections by Rosen, under contract to Transcanada Pipelines (TCPL), who have kindly provided recent results from inspections on their system[17-20].

TDW – SMFL

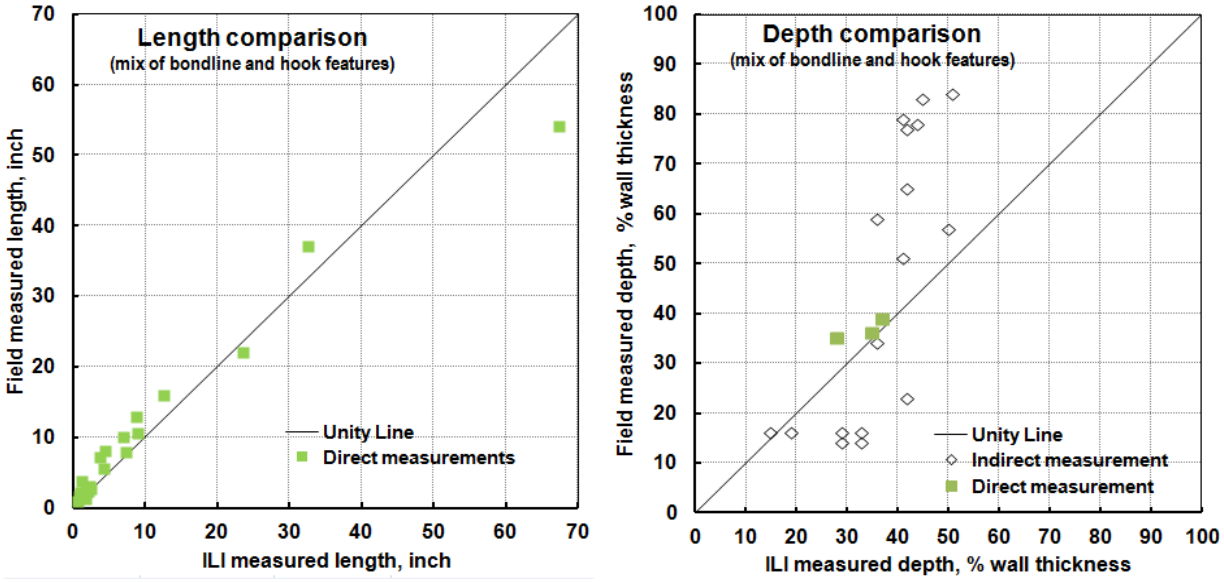
Background. As outlined in Reference 14, TDW has been running Multiple Data Set (MDS) ILI tools since 2009. Such combination tools are equipped with several inspection technologies, which for the SMFL tool couples the helical-field MFL with high-field and low-field axial MFL and ID/OD discrimination, high-resolution deformation, and an odometer – with inertial mapping available as an option. This combination of data sets is considered to be particularly useful in identifying very narrow but finite-width anomalies associated with pipe long seams. Anomalies for which the tool is applicable are quantified in geometrical terms in TDWs tool performance specification, which is included in Annex B as adapted to the circumstances considered in this reporting. The following presents a brief overview of recent seam assessment performance using this technology, as adapted from Reference 14.

In reference to Annex B, which includes the detection and sizing capabilities for the SMFL tool, it is clear that anomaly geometry is the basis to quantify detection and sizing in reference to ILI and other such tools. Reality in this context is that the ILI tool is not generally capable of precisely determining the type of anomaly relative to the usual categories (e.g., cold weld versus hook crack, etc.). For this reason, TDW chooses to use "planar anomaly" for axially oriented anomalies that manifests negligible width in the inspection data.

Data Trending. With this definition, and the notations in Annex B for the SMFL tool, the results of ILI runs can be presented as shown in Figure 13. The x-axis in this figure is the dimension called by the ILI, whereas the y-axis in the corresponding result based on the field measurements. Figure 13 presents data for a total of 21 anomalies that were reported as planar by the ILI. These results reflect six inspection runs, each for a different pipeline operator, for pipelines whose ODs ran from 8.625-inches up to 16-inches. Figure 13a shows the measured length data, while Figure 13b presents the measured depth data. Because the available data are limited, which precludes a broad statistical analysis, this comparison is limited to consideration of this unity plot.

Ideally, the measurements made in field digs provide a true estimate of the feature's size that is free of bias and embeds little uncertainty. Developing true bias-free results requires direct observation of the fully exposed feature and the use of tools whose resolution is fit for purpose and calibrated relative to traceable standards, such a digital microscope. Such measurements

provide a benchmark for the viability of the ILI results. However, as was noted above in regard to the data discussed in Reference 3, if the results developed in a field dig are based on other indirect measurements, such as those made using PAUT, they can be less reliable. Nevertheless, because PAUT remains the best ITD technology for such purposes, and also is convenient for use in such digs, this technology underlies the majority of the measurements reported on the y-axis in Figure 13. For this reason it is not a surprise that for the one case where use of PAUT was replicated, the independently measured depths reported by two UT crews were quite different. One crew reported the depth at 68% of wall, while the second crew reported it at 85% of wall. These very different depths are reported in Figure 13 as the average of these values, which is 76.5%.



a) length **b) depth**

Figure 13 Comparison of the SMFL measurement with reference data

In addition to the PAUT that was used for all features identified in the field as cold welds, three cases identified as hook cracks involved direct measurements developed using metallographic cross-sections. Of the depth measurements provided, the three directly measured depths can be considered benchmarks because they are traceable to calibrated standards and embed limited uncertainty. Although the depths quantified by PAUT are measured with a tool that is traceable to calibrated standards, as noted above indirect measurements based on commercially available tools embed uncertainty that is large relative to the features being sized. As such, these depth measurements are not considered as benchmarks. For this reason, Figure 13 considers results that are traceable to calibrated standards and embed limited uncertainty – as occurs for most direct measurements – to be benchmarked. It identifies such benchmarked data as solid square symbols, and discriminates these from indirect measurements – like those sized by PAUT – by representing those results as open diamond symbols, with both labeled accordingly. It is evident that all results shown in Figure 13a involve benchmarked lengths, whereas only three of the measured depths are considered benchmarks.

Results and Discussion. It is apparent from Figure 13a that the SMFL technology provides quite accurate sizing in regard to the benchmarked length measurements, for features that involve hook cracks as well as cold welds. Such results reflect cold welds in lengths from a little less than an inch to in excess of 50 inches, and hook cracks with lengths the order of an inch to two inches. The results suggest that typically the ILI-measured length slightly underestimates that of the benchmark value. In regard to Figure 13b, which presents the results for measured depths, it is also evident that depths quantified by the SMFL technology are quite close to the three available benchmark depths, which as noted above are specific to hook cracks. It is also evident that these three benchmarked results represent a narrow range of depths. Note in this context that while the limited results based on metallographic sections lie on the one-to-one trend, they also fit well into the scatter-band for the PAUT data. Until results that lie on the one-to-one trend are developed for a range of depths that fall well outside the PAUT scatter-band, little can be said that is conclusive.

A final point to note when comparing length and depth measurements is that anomalies have irregular shapes, with the depth varying over the length of the feature called or measured in the field much more so than the length. In spite of this, it is common for field crews to report a single length and a maximum depth for any anomaly. In contrast, the depth profile of typical axial planar features can be quite complex, which often show deeper central regions that trail off into very shallow tails at either end of the feature, where the feature falls below the detection threshold for the tool. For such anomalies the ILI length will generally under call the field length, as was evident in Figure 13a. Moreover, the vendor notes that amplitude of the helical-field (spiral) MFL signal response will generally quantify the effective depth of the entire anomaly profile, rather than the local maximum depth. Consequently, anomalies with short segments that are locally much deeper than the rest of the anomaly can lead to a depth under-call by ILI relative to field depth. Until all tools react similarly to the same feature traits, differences due to characterizing complex features using in a few simple measurements are inevitable.

In summary, the benchmarked results for length and depth appear to fit the one-to-one trend, but for the depth measurements also lie within a scatter band that runs diagonal to the one-to-one trend. Although that scatter reflects the nontrivial uncertainty in applications of PAUT to ERW seam defects[e.g.,3], until a broader benchmarked database is available the true utility of this MDS technology remains uncertain. In regard to anomaly classification, 20 of the 21 anomalies were correctly classified by the ILI tool, as both the cold welds and the shallow hook cracks are planar features. As such, the tool provided quite consistent discrimination of the features identified. The one feature misclassified as planar through-wall was a mid-wall lamination.

Combining multiple technologies in the context of the TDW MDS[15] analysis concept creates opportunities to compensate for individual technology limitations. For this application, SMFL can be used to detect and size planar axial anomalies because they produce a broad obstruction across the flux field. The absence of signal response in axial MFL, which is blind to such features, serves to confirm that a planar axial anomaly has been detected. In contrast, a response

to both axial MFL and spiral fields indicates that the anomaly is volumetric in nature. Therein lies the benefit and promise of the MDS concept.

Rosen – EMATs

Background. While Rosen also has developed and deploys a CMFL tool that can be used for axial flaw detection, this section considers their axial crack detection tool that relies on EMATs technology[9]. While this technology has been in commercial service since 2003[16], as for most technology it continues to undergo evaluation and refinement. Anomalies for which this tool is applicable are quantified in geometrical terms in Rosen's tool performance specification, which is included in Annex B after being adapted for present purposes.

TCPL has recently collaborated with Rosen and others in pull-testing and other analysis at Rosen's Technology and Research Center in Lingen Germany, which were focused on detection and sizing of anomalies in ERW seams using EMAT technology. The pipe was removed from a TCPL pipeline constructed of 22 inch diameter pipe with a 0.281-inch thick wall made of X-52 that was produced in 1956 by YS&T. YS&T production in 1956 is one of several years of their output that has been identified with a history of failures[21]. This pipeline was subject to a pre-service pressure test to near 72% of SMYS and subsequently hydro retested in 1970 to 96% of SMYS. Thereafter, there have been three seam-related leaks associated with cold welds the most recent of which in 2012 led to an ILI with the EMAT tool with supporting verification digs and a hydro retest to meet the US Code[13] required pressure (≥1.25 times the maximum allowable operating pressure (MAOP)).

Prior to the field inspection and subsequent hydrotesting, the EMAT tool was evaluated in pull-testing of pipe joints from the pipeline to be inspected into which twelve notches were electro-discharge machined (EDM) into the seam weld. This was done to create reference reflectors, which were clearly visible in the EMAT inspection. In contrast, shallow seam features assessed via UT shear-wave NDE as less than 2mm deep were only marginally evident in the EMAT inspection, as was one mid-wall inclusion. Destructive sizing via grinding indicated the depths of those cold-weld features were ~1.4mm and ~0.7mm (or 0.055 and 0.028 inch). Once this was preliminary testing completed, four sections of TCPL's pipeline that were located in high-consequence areas were run using the EMAT tool, after which the line was hydro retested.

A prior hydrotest had been done to 96% of SMYS coupled with TCPL's desire to establish the stability of any seam features that had survived the prior test meant that this retesting had a target maximum pressure the order of 96% of SMYS, or slightly higher. A modified spike test practice was used, with the hold time and leak-check pressure reduced in comparison to Code-required levels in consideration of suggested practices for lower toughness pipe[22]. This hydrostatic retesting resulted in two seam failures in one of the four retest sections, after which the target pressure was achieved.

Data Trending. Once the field ILI was completed, several joints were cut out for which the EMAT tool had shown indications in seven. These joints were sent to Lingen Germany for

detailed characterization by PAUT, and destructive testing – with aspects of this work still underway. As has been indicated in regard to the SMFL tool, successful anomaly sizing via any inspection practice is quantified in terms of the length and depth of the features detected relative to some benchmark measure of these parameters. As for that discussion, benchmarks to assess the viability of an ILI tool involve direct measurement at adequate resolution using tools that have been calibrated traceable to known standards. The present results rely on surface-measured lengths, along with lengths and depths measured via PAUT.

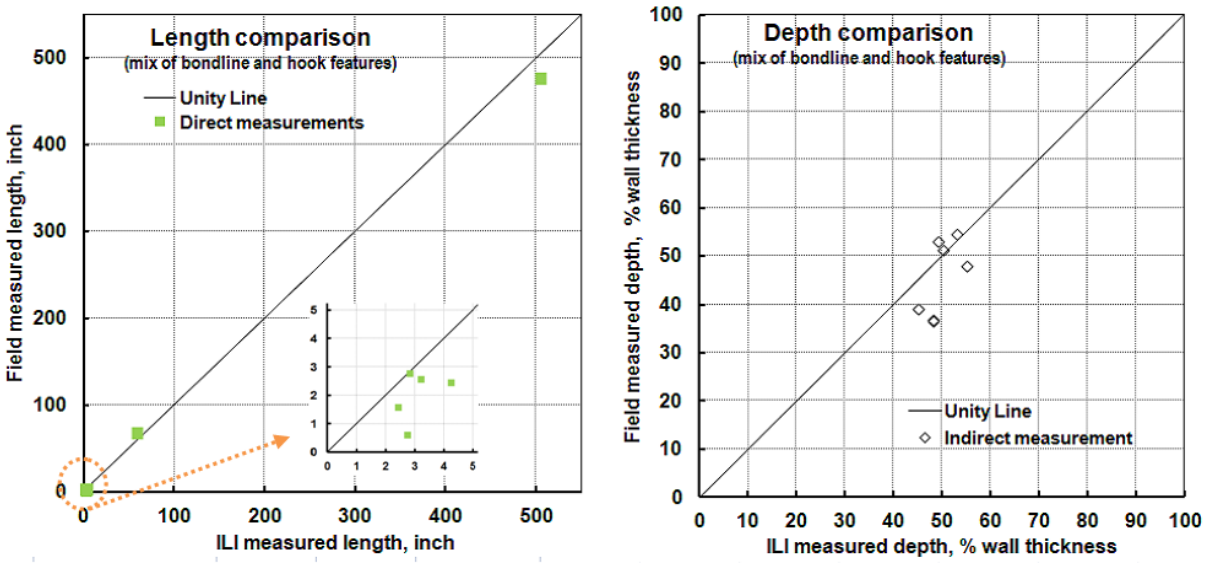
For surface-breaking features, the length often is longest at the surface, so visual measurements can be viable benchmarks when made relative to a standard steel rule. When the actual length is not clearly defined, the feature can be enhanced through use of black on white or other form of contrast-based magnetic particle inspection (MPI), such as wet fluorescent MPI, with this practice used selectively in subject study. For the present work, the directly measured lengths obtained from photographs that included a reference rule are a very close match to the PAUT length, which establishes their utility as benchmarks. However, the PAUT depths did not match the depths measured by grinding for the shallow features assessed as part of the EDM notch evaluation that was done prior to the ILI. For this reason, the indirect PAUT depth results are not considered benchmarks, whereas the direct and PAUT length measurements are. As these pipes continue to be destructively evaluated, results considered for the present to be benchmark values might change as other directly measured data are reported.

It is understood that like other inspection vendors, Rosen uses the term "planar anomaly" in regard to axially oriented anomalies that manifest negligible width in the inspection data. As for other inspection vendors, Annex B that includes the detection and sizing capabilities for the EMAT tool indicates that Rosen uses two parameters to quantify anomaly geometry. Thus, as for other inspection technologies, this opens to differences in what is apparently sensed on an area basis versus what is typically quantified by uncoupled measures of an overall length and a maximum depth.

As discussed in regard to the SMFL results, where reference dimensions for comparison to the ILI are measured with a tool that is traceable to calibrated standards and embeds limited uncertainty relative to the scale of the features measured, the result is considered a benchmark. As for TDW, TCPL's work in collaboration with Rosen and consultants made use of direct surface measurements to quantify anomaly length, and PAUT to quantify depth. For this reason, the same notation and symbol convention used in Figure 13 continues in use to present the results for the EMAT technology. Direct measurements are considered benchmarked and shown as solid square symbols, while indirect measurements – like those sized by PAUT – as shown as open diamond symbols, with both labeled accordingly.

With this background, the results of this collaborative study are presented in Figure 14, wherein the x-axis is the dimension called by the ILI, the y-axis in the corresponding field measurement. It is evident that all results shown in Figure 13a involve benchmarked lengths, whereas for these data all depth measurements are indirect. Figure 14a presents the measured length, while

Figure 14b presents measured depth. These figures show the results developed to date for the 7 anomalies reported by the EMAT ILI[17]. Because the available data are limited, this comparison is limited to consideration of a unity plot.



a) length **b) depth**

Figure 14 Comparison of the EMAT measurement with reference data

Results and Discussion. It is apparent from Figure 14a that the EMAT technology provides quite accurate sizing in regard to the direct length for longer features and PAUT depths for these cold welds. Because of the wide range of lengths measured, five of the seven results lie clustered together for the shorter features, which show that the ILI tool over estimates the lengths for shorter features (for results captured from photographs). Such results cover lengths from well under an inch to approaching the length of the seam in a typical double-random joint.

Whereas for the results presented for the SMFL tool showed it slightly underestimates length, the results for the EMAT tool indicate it slightly overestimates length. In regard to Figure 14b, which presents the results for measured depths, it apparent that the depths quantified by the EMAT technology closely match those measured by PAUT, but as for the SMFL tool the data represent a limited range of depths.

Other Collaborative Work

Other collaborative work directed at improved axial crack detection and sizing has been underway for several years involving Rosen, TCPL, and others[e.g.,17-20]. Since 2011, at least 12 pipeline segments with an aggregated length of more than 800 miles have been inspected with updated versions of EMAT tools. More than 80 excavations have been performed in support of this with direct examination done for the features identified to assess the viability of the inspection data. Retrospective analysis indicates that the bulk of the crack-like features evaluated were due to stress-corrosion cracking (SCC) – which while different from ERW seam cracking do involve tight planar reflectors and so pose a comparable challenge.

Analysis of these data show a significant improvement in tool performance has occurred over the past several years in regard to the probability of detection (POD) and the probability of correct identification[18] (POCI). For example, as quantified by the operator[18], over the interval from 2008 to 2011 the POD for cracks is above 90% at a 95% confidence level, which is well above the normally cited POD of 80% at the same confidence level. Recognizing that from an integrity perspective false positives are crack-like features that do pose a threat (such as laminations, narrow axial corrosion, and weld anomalies), their detection and call as cracks does not detract from the utility of the tool. In that context, the definition of POCI can be modified to the ratio of true positives to sum of true positives and false negatives, excluding the false positives. Performance quantified in terms of this modified POCI for lines with a statistically significant number of observations led to a success rate larger than 91% at 95% confidence level[18].

Analysis of these more recent data also show a significant improvement in tool performance over the past several years in regard to sizing depth and length[18]. As for the EMAT ILI depth and length sizing performance discussed in reference to Figure 14, the viability of this EMAT ILI was benchmarked by in-ditch NDE sizing referenced to PAUT, whose tolerance for that work was taken to be ±1mm. For purposes of the work being discussed, depth sizing was quantified in regard to the non-conservative ILI calls, which were considered to be out of tolerance, whereas the conservative depth calls were considered within tolerance. In that context, if the PAUT measured depth was within the ILI estimated maximum depth plus the depth error defined as (sqrt [ILI depth tolerance]^2 + [NDE depth tolerance]^2)[23], then it can be considered 'within specifications', otherwise it was 'outside specifications'. With this definition, the depth sizing success rate over the last several years was reported by the operator and its consultants to be in the range of 86% to 100%, with progressive improvements increasing the performance from 86% in 2008 up to 100% performance for the 2011 results.

Because of the nature of SCC colonies and differences in the reporting practices for a colony between the EMAT ILI and the ITD NDE report, sizing crack length is more difficult than sizing depth. For example, the ITD NDE often reports length based on the dimensions of the colony, whereas the ILI report is either as total length reported by the tool or an interlinked crack length based on ILI vendors' criteria. Accordingly, ILI performance in sizing length is assessed relative to both the interlinked length and the colony length as reported by the ITD NDE. In that context, if the absolute value of the difference in ILI and NDE measured length is greater than ILI length tolerance, the result is considered 'outside tolerance', otherwise it is taken to be 'within tolerance'. The success rate in sizing interlinked length according to the approach detailed in Reference 18 ranges from 0 to 33%. When only segments with a statistically significant number of observations are considered and sizing is referenced to colony length the success rate improved somewhat – but only marginally. While in view of these results length sizing is an area clearly in need of improvement, the situation is much better in regard to sizing length for features in an ERW seam – as evident in Figures 13 and 14.

Summary for Inspection

As noted in regard to the SMFL results, some error must be anticipated for any inspection tool when the irregular shape of an anomaly is simply quantified by a length and maximum depth. The key point to draw from the results in Figures 13 and 14 and the related discussion is that the more recently developed technologies or those involved in projects to refine or enhance their predictive capabilities are much improved in contrast to the outcomes reported in Reference 3. Such development and refinement is anticipated to continue through the efforts of Task 2 in the second phase of this project.

Implications for Pipeline Integrity Management

The observations of the prior section indicate that significant improvements in crack detection and sizing have been achieved in regard to the recently developed ILI tools, such as that based on EMAT technology. Such results are in contrast to the outcomes reported for the earlier generation technology as summarized in Reference 3.

It is apparent from the conclusions in Reference 3 that the utility of first-generation ILI and ITD inspection tools could be questioned as the basis for reliable condition assessment and integrity management. This supports the view of the National Transportation Safety Board (NTSB) made regarding such technology[24], as conveyed in correspondence to the Pipeline and Hazardous Material Safety Administration (PHMSA). There, the NTSB "concludes that current inspection … are not sufficiently reliable to identify features associated with longitudinal seam failures of ERW pipe prior to catastrophic failure" in such pipelines. They recommended[25] "that PHMSA conduct a comprehensive study to identify actions that can be implemented by pipeline operators to eliminate catastrophic longitudinal seam failures in ERW pipe" – which was the first of three related recommendations, and the genesis of the current project.

It is equally apparent that the PHMSA's belief "that in-line inspection technology is improving and data analysis capabilities are increasing each year" alluded to by the NTSB[24] is justified in view of the significant improvements evident in the work reviewed herein. Such results support the PHMSA's approach to integrity management, which relies on such condition assessment coupled with tools to quantify safe pressure limits and a timeline for re-validation to ensure continued safe operation through maintenance and rehabilitation on a fit-for-service basis.

Review of the viability of predictive tools, both in the broader context of Reference 2 and the more limited scope earlier in this report, indicates that the failure pressure can be reasonably predicted when the seam properties are understood, the sizes of the features are quantified and their location in seam is known. This observation supports the PHMSA's approach to integrity management – as the successful prediction of failure pressure is essential in its practical use. But, it is clear in view of References 2 and 5 that the local properties must be better understood and quantified, as they are central to viable predictions. Work is planned as a part of Task 4 of the Phase Two of this project to better quantify this aspect.

It is important to reflect on the failure pressures involved in the burst tests completed and their relationship to the sizes and locations of the anomalies causing failure as summarized in Table 1. Actual failure pressures the order of 96% SMYS as occurred for Pipe 16-16 are comparable to that used in field hydrotests. In contrast, failure pressures in excess of 130% SMYS as occurred for the other two pipes are well beyond anything practical under field conditions, as hydrotests pressures on the order of 115% caused deformations in the pipe that have been considered excessive in a Regulatory context[e.g.,26]. This work has identified features in the upset of the seam as well as in the bondline. In cases where both were present in the pipe, although failure occurred at quite high pressures, invariably a smaller bondline feature controlled that failure in spite of the presence of much larger features in the upset. Figures 6, 9, and 12 all show that anomalies that lie in the bondline pose a much greater threat than those in the upset, because at a given failure pressure they are smaller, and so harder to detect and size, and also statistically more frequent than the larger features. These figures indicate that the sizes of features in the bondline that can fail at in-service pressures are the order of an inch in length at maximum depths the order of half of the wall thickness, with shallower features failing at longer lengths.

The vendors ILI specifications for the SMFL[27] and EMAT[28] tools are summarized in Annex B after adapting them to consideration of the three pipe sizes used in the burst testing. These specifications indicate minimum detectable length and minimum detectable depths or depth accuracy, and so define the vendor's view of their tool's capability. These vendor specified sizes can be mapped into failure pressures through use of the trends in Figures 6, 9, and 12, to better understand the effectiveness of ILI as compared to hydrotesting, in light of the suggestion in Reference 18 that an ILI could be used in lieu of a hydrotest. Consider in this context the best of the detectable lengths and depths for the three pipes evaluated in reference to Annex B, which occurs for Pipe 22-11. When the parameters for that pipe as quantified in Annex B are evaluated in light of Figure 12, just detectable features are indicated to fail at about 90% SMYS, for the specific line pipe geometry involved. Thus, for this scenario ILI could be equivalent to a 90% hydrotest if it was 100% reliable – all else being equal. In contrast, the worst-case scenario that develops for Pipe 10-7 leads to the detectable features that fail at about 72% SMYS for the specific line pipe geometry involved. In this case, even if ILI were 100% reliable if would be much less effective in comparison to a hydrotest done to Code-minimum requirements.

One concern in view of this comparison is that the ILI equivalence that develops falls at or below code-minimum hydrotests levels, whereas for hydrotests to be broadly effective in exposing ERW seam defects history indicates[e.g.,6] that the maximum test pressure should exceed 100% SMYS. A second concern is that the effectiveness of detection as specified by the vendors is variable, depending on the pipe geometry and the properties within the ERW seam. The final arises from the observation that the ILI detection as it has been quantified by the vendors develops 80% confidence, whereas a pressure test will expose features that are critical at the test pressure. The downside to pressure tests from an integrity perspective is that they are applicable specifically to the instant in time that the peak test pressure is achieved. History demonstrates that another downside is that high-pressure tests can promote pressure reversals[e.g.,6]. Finally,

water must be obtained and then disposed of. It follows that there is no simple solution. Key in this context is that the minimum specified detection in terms of anomaly size does not translate into constant ILI effectiveness when viewed in terms of equivalent hydrotest pressure. Care must be taken in that context when more broadly assessing the implications of crack detection and sizing as presented in Reference 18.

At lower toughness levels as develop in some vintage ERW seams the failure boundaries as discussed in Figures 6, 9, and 12 lie at lower pressures as compared to situations where higher seam or pipe toughness can be relied on with certainty. For such cases, the improving capability of ILI suffices for reliable integrity management, and for some applications could replace hydrotesting for purposes of integrity verification. For very tough steels, defects large enough to grow in service cannot be exposed at practical hydrotest pressures[29], so there is a clear need for alternative integrity verification (AIV) tools based on inspection technology. In contrast, the toughness evident in some ERW seams leads to much smaller critical anomaly sizes, on the order of what can be reliably detected, as evident above. As inferred in some vendor specifications, detection of anomalies in seams also is more difficult. Thus, meeting the challenge to "eliminate catastrophic failures in ERW pipe; as well as to the vintage system"[24, 25] is much more demanding of the inspection technologies. It is more so an issue because the toughness is often uncertain. Accordingly, continued improvement in both ILI and ITD technologies is needed, as is a focus on correctly calling the type of feature and its location.

Summary and Conclusions

Results from full-scale burst-tests for electric-resistance weld seamed pipes with defects were presented and evaluated as the basis to assess the viability of available tools to 1) predict failure pressure and 2) to detect and size seam anomalies via ILI and ITD practices. Compare-contrast scenarios were developed for three burst tests relative to both inspection and failure prediction. The viability and reliability of these tools was discussed relative to ERW seam features and the implications more broadly assessed in regard to the vintage as well as the modern pipeline systems. Thus, the results of this work help to define actions essential to improve the integrity management practices for ERW seamed pipe, with the possibility that those outcomes will have implications for standards development, or tool development and commercialization.

Results were presented for the use recent of second or third generation technology to detect and size axial seam defects, specifically spiral magnetic-flux leakage (SMFL) and electromagnetic acoustic transducers (EMATs). Trending for these results showed that much improved detection and sizing can be achieved today as compared to the outcomes developed early in the use of the first-generation inspection technologies. Trials using emerging ITD technology referred to as Inverse Wave Field Extrapolation (IWEX), which couples phased-array ultrasonic technology (PAUT) and time of flight diffraction (TOFD), were promising. With such technology indicated that step improvements are possible compared to currently available tools. That being said, given the limitations with the ITD technologies currently available the most reliable approach couples magnetic particle inspection with TOFD and PAUT.

The defects causing failure were found in the bondline, as well as in the upset region of the seams considered, which trace back to manufacturing setup and process upsets. Results developed showed that anomalies in the upset of the seam were much more stable than those in the bondline, which made clear that size alone does not define the threat posed by an anomaly. Metallography and fractography made clear the complexity of real seam defects, as compared to machined (idealized) features, not only in regard to their shapes and sizes but also in regard to microstructural differences that can affect failure response. It follows that there is a need to identify the location as well as the type of anomaly if such features are to be prioritized in condition assessment following the inspection.

It was found that reasonable predictions of failure pressure were possible for ERW seams when the shapes and sizes of the features were known, and the toughness local to the failure site could be estimated based on local properties data. This means that the models used to quantify failure pressure must be specific to the type of defect: that is bondline versus hook crack versus selective seam corrosion. While good predictions could be achieved when the differences in the severity of the features, and the local resistance to failure were addressed, scatter was evident when more rudimentary analyses were done based on nominal properties. As such, uncertainty in local toughness and UTS can cause scatter in the predicted failure pressure, as can inadequate anomaly sizing.

Useful conclusions can be found throughout the report, of which the most important follow:

- ILI done using SMFL and EMAT tools focused in part on crack-like features associated with SCC over almost 1500 miles of liquid, highly volatile liquid, and natural gas pipelines made using low as well as high frequency ERW processes showed the technology to detect cracking has recently improved significantly. Based on data reported by the operator and their vendors –
  - over the interval from 2008 to 2011 the probability of detection (POD) via EMATS for cracking due largely to SCC was found to be above 90% at a 95% confidence level, which is well above the normally cited POD of 80% at the same confidence level;
  - in contrast to failures on recently inspected lines using earlier generation technology, results specific to recent EMATs technology indicate that the probability of correct identification for lines with a statistically significant number of observations led to a success rate larger than 91% at 95% confidence level.
  - likewise, in contrast to failures on recently inspected lines using earlier generation technology, results specific to recent EMATs technology indicate that the success rate for probability of correct the depth sizing has shown progressive improvements from 86% in 2008 up to 100% in 2011.
  - because these results are in strong contrast to past experience and the expectations of some experts, there is a need to better understand and document the circumstances that underlie the improvements and more broadly replicate these observations.

- collaboration between vendors and operators, and experts as needed, has contributed greatly to the improved detection and sizing capabilities;
- vendor specifications for ILI tools were found in some cases to be equivalent to a 90% SMYS hydrotest, but this outcome was confined to specific combinations of line-pipe geometries, as for some geometries the tools were indicated to be less effective;
- the means to establish the viability of an ILI run via ITD technologies like phased array ultrasonic technology can be less reliable than desired;
- limited testing with emerging ITD tools based on PAUT indicated step improvements in anomaly sizing will be evident as compared to the status quo once such technology becomes commercially available;
- the irregular shape of real anomalies makes it difficult to quantify their size using the usual two parameters – maximum depth and length – which confounds assessing the viability of ILI;
- differences between the measurements from different sensor technologies are inevitable so long as complex features are characterized using in a few simple measurements – which also confounds assessing the viability of ILI;
- failure is controlled by feature size and also the local properties, such that the interpretation of ILI and ITD tools must be taken in light of the features location, and the properties to the extent they can be inferred – likewise the development of inspection tools to routinely quantify local strength and toughness would affect a step improvement in failure pressure predictions;
- meeting the challenge to "eliminate catastrophic failures in ERW pipe; as well as to the vintage system" is demanding, with continued improvement in both ILI and ITD technologies needed, including a focus on correctly calling the type of feature and its location – in addition to detecting and sizing it: and
- while the claim was made that inspection could be used to replace hydrotesting, and some success was noted that supports this view, it was also evident that in some cases the ILI equivalence that developed fell at or below code-minimum hydrotests levels – in addition it was evident that the effectiveness of detection as specified by the vendors is variable, depending on the pipe geometry and the properties within the ERW seam: key in this context is that the minimum specified detection in terms of anomaly size does not translate into constant ILI effectiveness when viewed in terms of equivalent hydrotest pressure.

References

1. Nestleroth, J. B., Smith, J. M., and Leis, B. N., "Data on Pipe Collected: Inspection by in the Ditch Methods, In-Line Tools, and Hydrostatic Tests," Battelle Interim Report, Subtask 2.1/2.2, US DoT Contract No. DTPH56-11-T-000003, May 31, 2013.
2. Kiefner, J. F. and Kolovitch, K. M., "Models for Predicting Failure Stress Levels for Defects Affecting ERW and Flash-Welded Seams," KAI Interim Report on Task 2.4, US DoT Contract No. DTPH56-11-T-000003, January, 2013.
3. Kiefner, J. F. Kolovich, K. M. Macrory-Dalton, C. J., Johnston, D. C., Maier, C. J., and Beavers, J. A., "Track Record of In-Line Inspection as a Means of ERW Seam Integrity Assessment," KAI Interim Report on Task 1.3, US DoT Contract No. DTPH56-11-T-000003, November, 2012.
4. G. T., Quickel "Characterization of the Toughness of Pipe Containing Seam Defects," DNV Interim Report on Task 2.3, US DoT Contract No. DTPH56-11-T-000003, February, 2013.
5. Leis, B. N. and Nestleroth, J. B., "Characterizing Aspects of Failure in ERW Line Pipe," Battelle Interim Report, Subtask 2.6, US DoT Contract No. DTPH56-11-T-000003, April 25, 2013.
6. Leis, B. N. and Nestleroth, J. B., "Battelle's Experience with ERW and Flash Weld Seam Failures: Causes and Implications," Battelle Interim Report, Subtask 1.4, US DoT Contract No. DTPH56-11-T-000003, September 20, 2013.
7. Leis, B. N., "Time-Trending and Like-Similar Analysis for ERW-Seam Failures," Battelle Interim Report, Subtask 4.2, US DoT Contract No. DTPH56-11-T-000003, April 15, 2013.
8. http://www.tdwilliamson.com/en/Services/PipelineIntegrity/spiral_mfl/Pages/Home.aspx
9. http://www.roseninspection.net/RosenInternet/InspectionServices/ILInspection/CrackCoatingDetection/
10. Leis, B. N., "Pipeline Axial Flaw Failure Criterion, PAFFC," Release 4.5, delivered in conjunction with Leis, B. N., and Eiber, R. J., "Fracture Control Technology For Natural Gas Pipelines – Circa 2010", PRCI Report Mat4-5, 2013.
11. Katz, D. Private correspondence, 28 and 29 May, 2013
12. Anon., "Specification for Line Pipe," API Specification 5L, American Petroleum Institute, Dallas, 43rd Edition.
13. US Federal Pipeline Safety Regulations 49 CFR Part 192, Subpart J
14. Ludlow, J., "Summary of Recent Seam Assessment Performance using TDW Multiple Data Set Tools," March 2013.
15. Harris, C., "Summary of Recent Seam Assessment Performance using TDW Multiple DataSet Tools," Pipeline and Gas Journal, March, 2011.
16. Beuker, T., and Brown, B., "The Evolution of an In-Line Inspection Solution: Axial Flaw Detection," http://www.roseninspection.net/RosenInternet/MediaArchive/Presentations/
17. Weber, R., "EMATS Data Analysis for TCPL's ANR 504 Pipeline," April, 2013.
18. Gao, M., Tandon, S., Krishnamurthy, R., Kania, R., Rosca, G., and Piazza, M, "A Comprehensive Analysis of EMAT ILI and In-Ditch SCC Inspection Technologies and their Application to Pipeline Integrity Management Programs," Paper 20, 19th Biennial Joint Technical Meeting, Sydney, May 2013.
19. Gao, M., Tandon, S., Krishnamurthy, R., "Evaluation of EMAT Tool Performance by

Monitoring Industry Experience", Paper 7, 18th Biennial Joint Technical Meeting, San Francisco, May 2011.

20. Marr, J., Riverol, E., Weislogel, J., Tappert, S., Mann, A., Sun, J. "Validation of EMAT In-Line Inspection Technology For SCC Management", Proceedings of the 2010 International Pipeline Conference, Paper No: IPC2010-31091, Calgary, 2010.
21. Clark, E. B., Leis, B. N., and Eiber, R. J., "Integrity of Vintage Pipelines," Battelle Final Report including seven annexes to INGAA Foundation, October 2004 (http://www.ingaa.org/cms/31/7306/43/678/6143.aspx).
22. Leis, B. N., Galliher, R. D., Sutherby, R. L., and Sahney, R., "Hydrotest Protocol for Applications Involving Lower-Toughness Steels," *International Pipeline Conference, IPC04-0665*, Calgary, October 2004, pp.
23. Anon., API 1163, "Inline Inspection System Qualification Standard", American Petroleum Institute, 2005.
24. Anon., NTSB Safety Recommendation P-09-1, Letter Commentary dated October 27, 2009 to the PHMSA
25. Anon., http://www.ntsb.gov/news/events/2009/carmichael_ms/index.html, NTIS Number: PB2009-916501, Adopted October 14, 2009
26. Leis, B. N., Kurth, R. E., Zhu, X-K., and Clark, E. B., "Fitness-for-Purpose for the 36-inch Diameter Segments of the Boardwalk System," Final Battelle Report to Boardwalk Pipeline Partners, LP, June 1, 2010.
27. http://www.tdwilliamson.com/en/Services/PipelineIntegrity/spiral_mfl/Pages/Home.aspx; then click the pdf link to SpirALL® Magnetic Flux Leakage Technology Flyer
28. http://www.roseninspection.net/RosenInternet/InspectionServices/ILInspection/CrackCoatingDetection/ then click the link to RoCD^2 Flyer (pdf)
29. Leis, B. N., Appendix B to "Hydrostatic Testing Of Transmission Pipelines: When It Is Beneficial and Alternatives When It Is Not", PRCI Catalog No. L51844, 2001.

Annex A: Details of the Pipes Evaluated

Annex A: Details of the Pipes Evaluated

This Annex presents the properties for the three steels considered in the burst testing. Each of the pipe steels is characterized in regard to its mechanical properties via tensile test data, as well as its hardness, chemistry, microstructure, and fracture resistance via CVN data. All testing was done in accordance with relevant ASTM or API Standards. To facilitate comparison between the three steels, the results for all three steels are grouped and discussed by property.

Tensile Response and Hardness

FigureA1 characterizes the mechanical properties of the three line pipes considered in this project. As outlined in Reference 1, all steels were specified as Grade X52. The pipelines that Pipes 16-16 and 22-11 were removed from were built in the mid-1950s, while that for Pipe 10-7 was built in the early 1960s. All seams appear to be using a LF process. It is apparent from the details embedded within Figure A1 that the pipes are consistent with the specifications for Grade X52 in regard to the aspects noted. Two of the steels behave more or less typical of this grade, although the actual yield stress (AYS) is higher than usual for steels of that production era. The third steel for Pipe 10-7 has a much higher UTS than is usual, in addition to having a value of AYS this is comparable to the other pipes. The values of the Brinell hardness for these pipes are quite typical of the era produced.

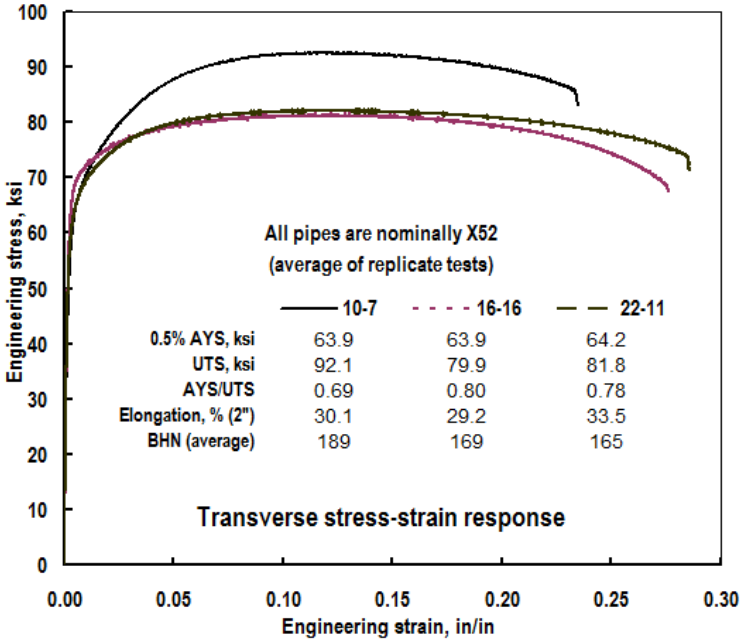


Figure A1. Mechanical and related properties for the pipes evaluated

Chemistry and Microstructure

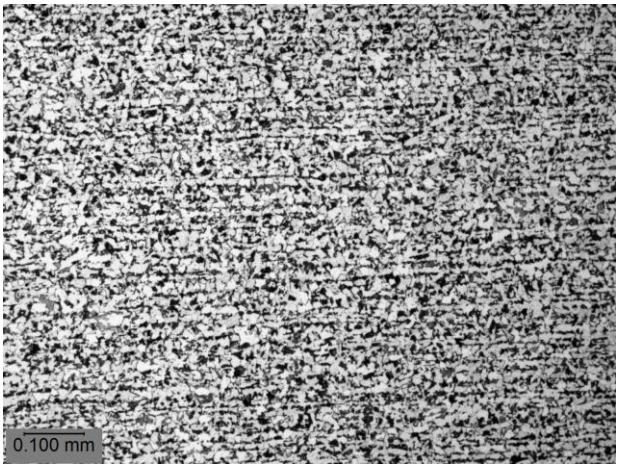
The chemistries for the three pipes evaluated are presented in Table A1 in regard to the elements specified in the era these steels were produced. Comparison of these results with the levels

specified for Grade X52 indicates that these pipes satisfy the grade requirements for electric-welded (ERW) pipe in the era they were produced.

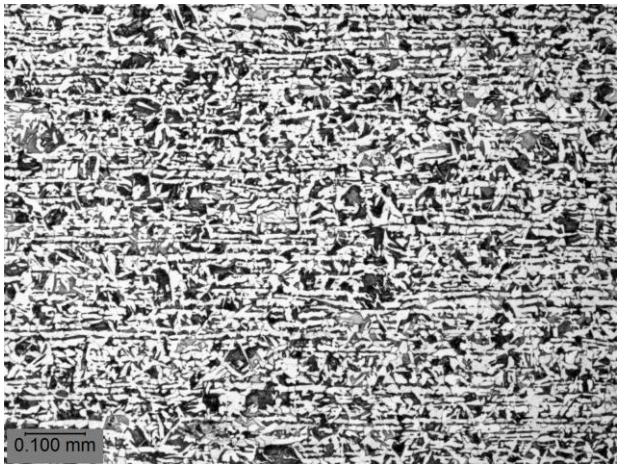
Table A1 Chemistries of the steels, in weight-percent

Steel	Composition in wt% of elements specified in the 1950s timeframe			
	C	Mn	P	S
10-7	0.279	1.090	0.009	0.021
16-16	0.262	0.780	0.013	0.032
22-11	0.250	1.140	0.014	0.021

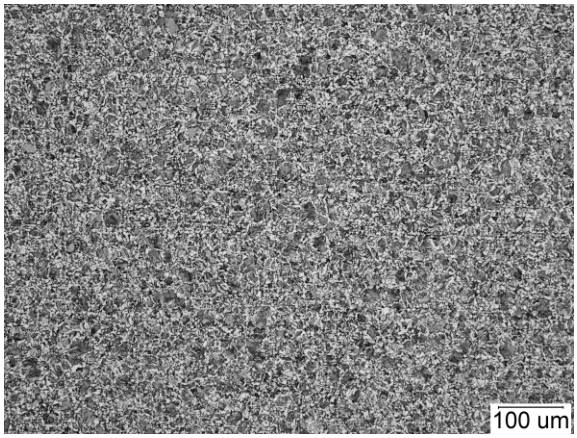
Figure A2 presents views of the microstructures for the three steels evaluated. It is evident from these images that the structure for two of the steels is a banded mixture of ferrite and pearlite, which is typical of the microstructures observed in hot rolled pipe steels for the chemistries listed in Table A1. The structure for 10-7 was much less banded, but still a mix of ferrite and pearlite.



a) 16-16



b) 22-11

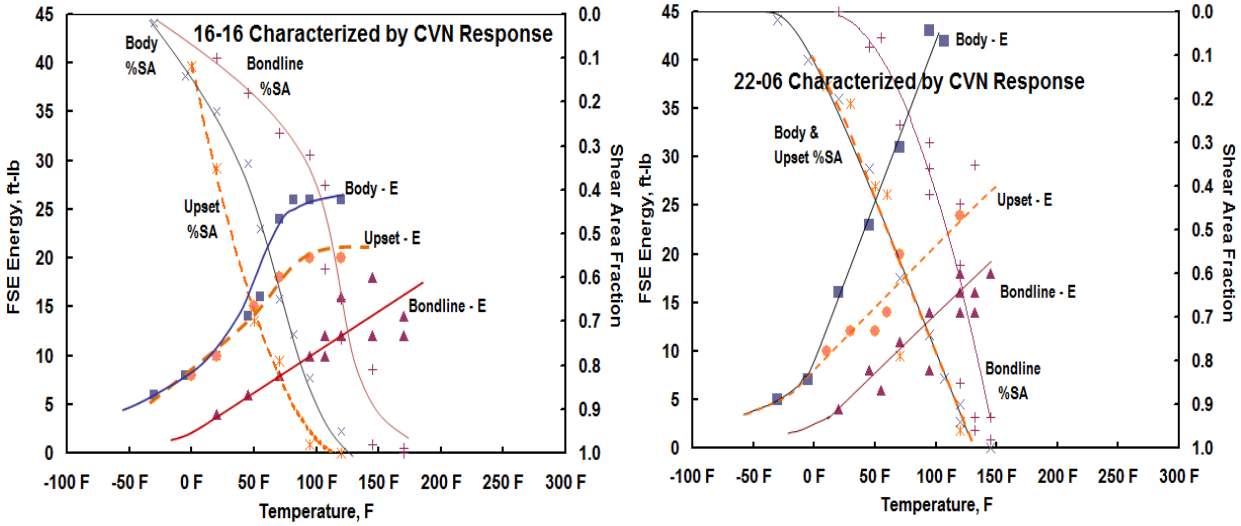


c) 10-7

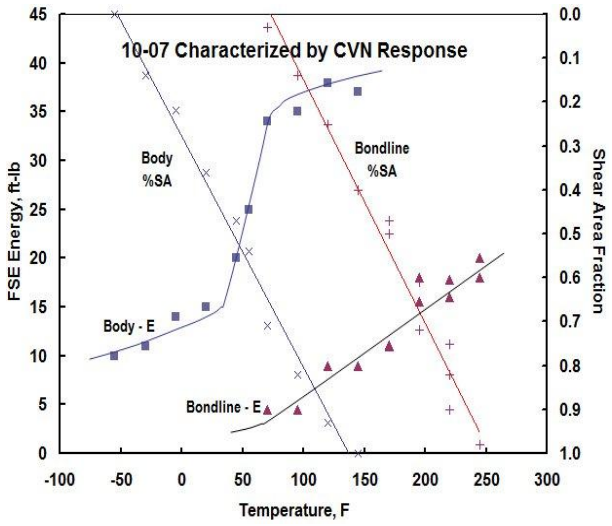
Figure A2. Microstructures for the line-pipe steels evaluated

CVN Resistance

The results of CVN testing for transverse samples notched in the through wall are shown in FigureA3 for samples cut from the pipe body as well as the ERW seam. Samples in the seam were notched to initiate cracking in the bondline or notched slightly off the bondline with the determination of this position verified after testing by etching the broken sample.



a) 16-16 b) 22-11



c) 10-7

Figure A3. CVN results for the pipes evaluated

While fractures were found to run in both the bondline and the upset for Pipes 16-16 and 22-11, such was not possible for the seam in 10-7. Results for these data sets are used specific to the test temperature of the burst testing in predicting the failure pressures for the defects that caused failure in each burst test.

Annex B: In-Line Tool Inspection Specifications

Tables presented in this annex reflect the vendor's published specifications for the SMFL and EMAT tools[27,28] adapted specific to the wall thicknesses for the pipes evaluated.

T D Williamson

Axial and Spiral MFL Combined Data Sets Adapted for ERW Crack Applications

	< Axial Slotting[1] >			< Axial Slotting[2] (most relevent W/A) >		
	Pipe 16-16: t = 0.252"	Pipe 10-7: t = 0.231"	Pipe 22-11: t = 0.290"	Pipe 16-16: t = 0.252"	Pipe 10-7: t = 0.231"	Pipe 22-11: t = 0.290"
Depth as d/t at POD = 90%	0.200	0.200	0.200	0.250	0.250	0.250
Depth accuracy as d/t @80% confidence	±0.20	±0.20	±0.20	±0.250	±0.250	±0.250
Width accuracy @80% confidence	±0.394 inch			±0.394 inch		
Length accuracy @80% confidence	±0.787 inch			±0.787 inch		

Limitations not addressed above:

In the vicinity of seam welds or heat affected zones, depth accuracies degrade by 0.10t and detection thresholds increase by 0.10t.

Detection thresholds and inspection accuracies in bends and other fittings are unspecified.

Notes:

1 Axial slotting in range L/A>=2, W/A>=0.5 — Implies they don't claim to size tight cracks: minimum length = 0.788" herein

2 Axial slotting in range L/A>=5, W/A<0.5, W>=0.010" — Implies they don't claim to size tight cracks: minimum length = 1.97" herein

Where 'L' and 'W' are actual Length and Width of the anomaly and the geometrical parameter 'A' is linked to the NDE methods as follows:

If t < 0.394" then A = 0.394", whereas if t >= 0.394" then A = 't'

Other requirements apply that pertain to operations and other factors - see NACE SP0102-2010

Rosen Performance Specifications for RoCD2 Tool

	Pipe 16-16 t = 0.252"	Pipe 10-7 t = 0.231"	Pipe 22-11 t = 0.290"
Crack Detection			
Minimum depth (parent material) = 0.04" as d/t	0.159	0.173	0.138
Minimum depth (in long seam) = 0.08" as d/t	0.317	0.346	0.276
Minimum length = 1.57"	<	1.57"	>
Crack sizing			
Depth sizing (±) = 0.15t as d/t	±0.150	±0.252	±0.231
Length sizing (±) = 0.39"	<	0.39"	>
Notes:			
Wall thickness range up to 0.79"			
Orientation to pipe axis is within ±18°			
Other requirements apply that pertain to operations and other factors - see NACE SP0102-2010			