Proceedings of the 8th International Pipeline Conference
IPC2010
September 27-October 1, 2010, Calgary, Alberta, Canada

IPC2010-31419

BALANCING PIPELINE SAFETY AND COST INTEGRITY MANAGEMENT THROUGH PERFORMANCE VALIDATION OF IN-LINE INSPECTION DATA

Rick McNealy
Applus-RTD
Houston, TX

Sergio Limon-Tapia
Williams Gas Pipeline
Salt Lake City, UT

Richard Kania
TransCanada.
Calgary, AB

Martin Fingerhut
Applus-RTD
Houston, TX

Harvey Haines
Kiefner & Associates, Inc.
Vienna, VA

ABSTRACT

In-Line Inspection (ILI) surveys are widely employed to identify potential threats by capturing changes in pipe condition such as metal loss, caused by corrosion. The better the performance and interpretation of these survey data, the higher the reliability of being able to predict the actual condition of the pipe and required remediation. Each ILI survey has a certain level of conservatism from the assessment equations such as B31G and sensitivity to ILI performance for measurement uncertainty. Multiple levels of conservatism intended to limit the possibility of a non-conservative assessment can result in a significant economic penalty and excessive digs without improving safety. A study was undertaken to evaluate the reliability of responses to ILI corrosion features through multiple case studies examining the effects of failure criteria and data analysis parameters. This paper discusses the effect of validated ILI performance on safety, and addresses the risk of false acceptance of corrosion indications at a prescribed safety factor. The cost of unnecessary excavations due to falsely rejecting ILI predictions is also discussed.

INTRODUCTION

Pipelines are designed and constructed recognizing prescribed factors of safety. These safety factors must be maintained through the life of pipelines. In-Line Inspection (ILI) is often employed to understand the condition of pipelines with respect to threats that cause metal loss, deformation or cracks. The ASME standard B31.8S: Managing System Integrity of Gas Pipelines advises that the results of ILI only provide indications of defects, with some characterization of the defects and that screening of the ILI information is required in order to determine the time frame for examination and evaluation.[1] It has been well established that the normal expected defect sizing performance of magnetic flux leakage MFL tools provides acceptable levels of safety equivalent or better than those established by hydrostatic testing. Measures of ILI performance are obtained from either laboratory type pull tests on a known, often fabricated, defect population, or from NDE field verification measurements using direct examination. ILI systems are employed recognizing tool error from published vendor performance specifications for detection and sizing performance. However, it has also been established that ILI tool performance can differ from the claimed performance due to a number of factors including the shape of actual defects, magnetite in the corrosion pits, and other run conditions.

An example of validated depth performance outside of claimed performance is shown in Figure 1. This data (from the Kiefner & Associates, Inc. database of ILI surveys) shows non-conservative MFL predictions where narrow pinhole corrosion was the likely cause. For ILI tools with performance that is consistent with the claimed specifications, there will always be a potential for some non-conservative predictions.

The relative conservatism among defect assessment criteria ASME B31G, modified B31G, and Effective Area Assessment (EAA) methods is well understood.[2] [3] Using ILI indications with the most conservative assessment (e.g., B31G), does not necessarily guarantee safety and could adversely impact timely response because of a large number of excavations required for direct examination.

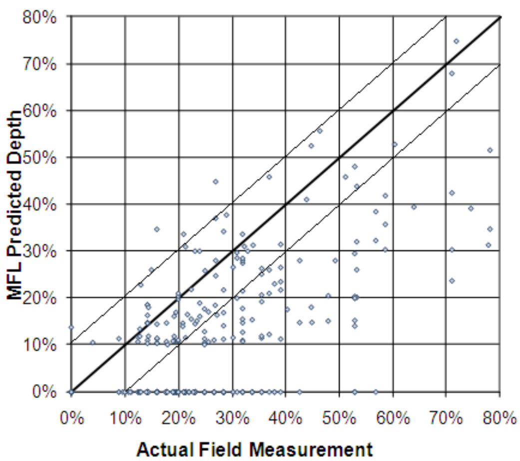

Figure 1: Validated Depth Performance from MFL, Pinhole Stray Current Corrosion

Based on measurement uncertainty, ILI predictions have likelihood for false acceptance (under conservatism) of a defect, which is a safety consideration. There is an alternate likelihood associated with ILI predictions regarding false rejection which results in unnecessary excavations and recoating with the resultant economic costs. Additionally, false rejections (over conservatism) can impact safety due to the potential for incorrect response prioritization on a program basis and numerous unnecessary excavations can potentially lead to damage of the pipe. Confidence in ILI predictions requires an understanding of true ILI tool performance. Consideration for ILI performance is also a regulatory requirement in some jurisdictions.

Not all pipelines are in a condition where integrity is sensitive to inspection performance. API 1163 provides a process for verification that reported inspection results are consistent with the performance specification identifying two possible outcomes [4]:

- Inspection Results Verified: Accept the ILI predictions without field verification, based on prior performance data
- Field Verification Measurements Recommended: Validate the performance claims
    - Accept or reject claimed performance

Mature integrity management programs that have progressed past the initial baseline assessments and remediation, into the re-inspection phase, can expect the ILI assessments to predict fewer severe defects which require immediate or scheduled remediation compared with the baseline results. A similar result can be expected from base line inspections of newly constructed pipeline segments. The API 1163 process recognizes that, in the event ILI does not report features large enough to be successfully located for field measurement, the consideration of other historical condition data and comparison of the ILI results with prior validated inspections from the same inspection system on other lines, supplemented with data from large scale tests, could be used to accept the inspection results. API 1163 fails to define a "large-scale test" but within the context of API 1163 as a whole, the concept of sample plans with sufficient population to provide 95% confidence suggests an approach to justify ILI system performance. Pooled validation databases or direct examination methods known to capture statistically large data sets offer options for documenting ILI system performance.

In many instances ILI integrity assessments predict large and severe defect populations. One consideration for a response plan could be to minimize the likelihood of false acceptance and false rejection of ILI predictions. To achieve this objective, the value of higher accuracy defect assessment criteria applied to ILI such as B31G Modified (0.85dL) or Effective Area Assessment (i.e., API 579/ASME FFS-1, RSTRENG) and others compared with ASME B31G have been well documented. However, the use of higher accuracy methods can be sensitive to the accuracy of ILI predictions. In the presence of a validated corrosion threat and sensitivity of immediate and future integrity to the defect population distribution, API 1163 recommends field measurements to validate performance.

SAFETY- MEASUREMENT UNCERTAINTY

ASME B31.8S recognizes the inherent conservatism associated with B31G assessment, and recognition for corrosion growth within its response schedule for features with safety factors less than class location requirements predicted by ILI. B31G assessment results are likely to reflect useful conservatism over a significant range of measurement error for feature depth and length. Therefore the effects of ILI performance are rarely considered within the application of the B31G criterion. The alternate assessment criterion B31G Modified (0.85dL) and Effective Area Assessment methodology are known to be fundamentally more accurate and less conservative than B31G. The potential for non-conservatism, related to ILI error, adversely affecting pipeline integrity was evaluated within the context of the following case study.

A test of ILI depth and length uncertainty and the resultant error in predicted burst pressure was conducted as a case study for a 20 inch natural gas pipeline, X-52, 0.219 nominal wall with 820 psi MAOP. To evaluate the effect of both depth and length random errors acting independently on burst pressure, a Monte Carlo type method was applied. Corrosion growth was also considered by incrementing the nominal defect dimensions at a rate of 5 mils per year. The results (summarized in Figure 2) were expressed as the likelihood for burst pressure to be less than MAOP for an evaluated metal loss feature (40% to 48% wall thickness depth, and 2 to 3 inches long) accepted at a prescribed safety level after 10 years. Figure 2 shows that B31G predictions can provide useful protection against unintended failure over a wide range of both depth and length measurement error. This level of conservatism comes at a price; B31G assessments risk identification of a large population of ILI response locations that will not require repair after evaluation performed by direct examination using EAA.

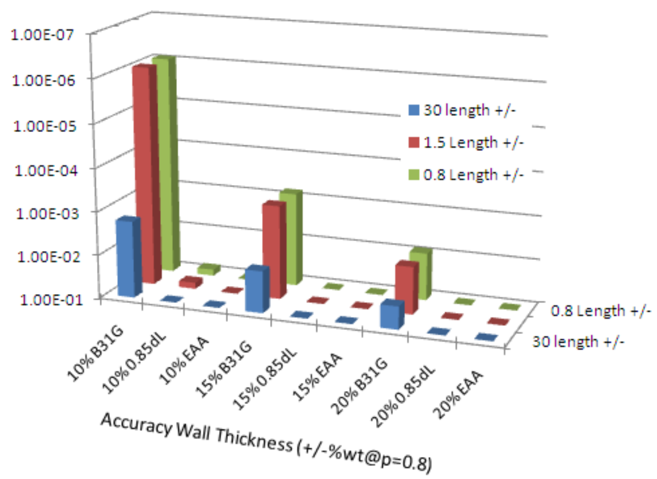

Figure 2: Predictions of likelihood (p) for rupture of an initial defect accepted at SF=1.39 over 10 years as a function of initial measurement error.

Effective management of large defect populations drives the need to manage conservatism through increased accuracy using 0.85dL or EAA. Rather than the extended response schedule embodied in ASME B31.8S, the alternate criteria offer adequate protection against false acceptance error with limitations based on ILI measurement error. In fact, other research has theoretically demonstrated the EAA method has a smaller uncertainty for longer multi-pit corrosion anomalies (less than one-half the uncertainty of the 0.85dL method when assessing a long enough defect with many depth measurements along the length of the defect). [5]

The value for applying EAA to ILI data lies in reducing the inherent conservatism compared with the other criteria and taking advantage of reduced uncertainty to insure safety. Figure 3 illustrates the effect of ILI sizing performance on the likelihood for false acceptance. The same case described for Figure 1 is used except the response time interval is now one (1) year. False acceptance (non-conservative treatment of metal loss features as safe) is reasonably unlikely ($p<1 \times 10^{-4}$) for both 0.85dL and EAA assessment when the ILI depth sizing performance is better than ±15%wt with 80% certainty and feature length predictions are within ± 1.5 inches. For assessments where a high accuracy criterion is applied and ILI indications near the repair limit would be accepted as safe, the value inherent in considering tool performance is clear. Remedial response may then be employed as an option to achieve an increased safety factor in the event the re-stated ILI performance is insufficient to provide a confident prediction of pipeline condition. Techniques for calibrating ILI predictions (the log) based on validated performance have been proposed by others. The reliable application of calibration techniques is dependent upon knowing the "true" performance of the ILI system which requires consideration of errors associated with the field validation measurements. [6]

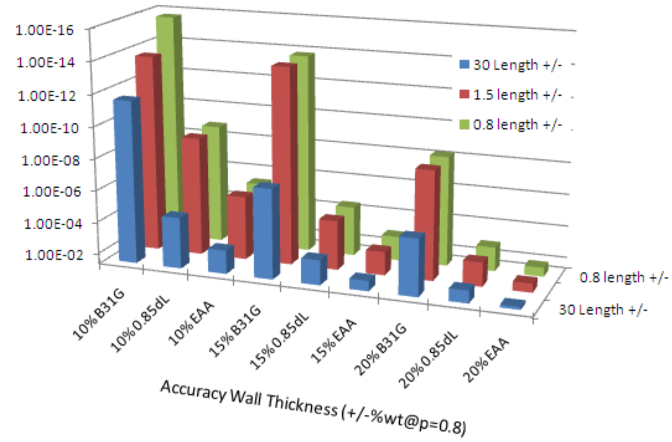

Figure 3: Predictions of likelihood (p) for rupture of an initial defect accepted at SF=1.39 in one year as a function of initial measurement error.

VALIDATION PERFORMANCE

Accuracy of the base sensor is but one component of the total systemic or process capability for an inspection technology. For MFL technologies the signal analysis process is another component. Validated ILI performance based on field measurements reflects the influence of direct examination system errors inherent in the techniques employed: e.g., ultrasonic gauge, laser profilometry or pit gauges. [4]

A wide variety of sensors and techniques are available to perform validation measurements. The systemic performance of pit gauges can vary from ±5%wt to ±10%wt when comparing measurement of independent, small diameter pits versus larger diameter, round bottomed pits where repeatability is difficult. Ultrasonic thickness measurement of small diameter flaws is limited by transducer frequency. Various transducer frequencies ranging from 2 to 10 MHz are routinely applied with high frequencies providing the ability to detect smaller diameter features, but size sensitivity (accuracy) considering beam spread must be taken into account. The thickness measurement for ultrasonic technology represents the most accurate validation method available (± 0.01mm), but total process capability can be affected by the mode of mechanical application. Laser systems depend on the reflected energy and for the pit diameter limit (1mm) there is a limit to the depth of a straight wall feature beyond which there is insufficient reflected laser light for measurement. The claimed performance for laser depth measurement in pipeline applications has been shown to be ±2%wt at p=0.8. [7]

Locating and matching ILI prediction locations to direct examination location is another potential source of systemic error to consider. Various techniques for locating and matching validation locations and their characteristic uncertainty affect the total error. [8]

What level of direct examination technology is required to support ILI validation? The answer depends on the sensitivity of pipeline integrity to both ILI prediction error and the size tolerance of the validation effort. When applying inherently conservative criterion such as ASME B31G, the need for high confidence in understanding true performance may not be so important. However, when there is a need to apply more accurate, less conservative criteria (EAA) the highest accuracy validation techniques (UT and Laser) can insure that pipeline integrity is evaluated against the true ILI performance, not the "apparent" validated performance.

ECONOMICS AND PROGRAM EFFECTIVENESS

The need for accuracy and high confidence in understanding ILI performance depends on the sensitivity of pipeline integrity to measurement error and the size of the task considering the need for prompt response to critical ILI predictions. As previously discussed, there are situations where it makes sense to use the highest accuracy to minimize costs but this approach requires recognition of the possible errors in order to assure safety.

<u>Case Study 1</u>

Table 1: Case Study Integrity Assessment Response Options

		# Features w/ SF< 1.1	# Features w/ SF< 1.39	# Joints w/ SF< 1.39
Pipeline Diameter, inch	26			
Wall Thickness, inch	0.281			
Grade (SMYS)	X-52			
MAOP, psig	809			
Design Factor	0.72			
ASME B31G		284	2427	> 500
B31G Mod (0.85dL)		99	858	393
EAA		1	100	63

A significant portion of the overall budget for ILI based integrity assessments likely consists of costs for excavation, direct examination and remedial action consisting of coating or pipe repairs. While the value of RSTRENG has been demonstrated to minimize unnecessary repairs, the potential for cost savings is often diminished by excavation costs incurred based on ILI metal loss characterization. As discussed earlier, the application of EAA to ILI is sensitive to the true ILI measurement error in order to preserve the superior uncertainty inherent with EAA compared with B31G or B31G modified (0.85dL). The difference in predicted repairs by ILI compared with field results validated by RSTRENG can be significant.

The following case study illustrates the value for high accuracy assessment of ILI data and the impact of tool performance on the results. The baseline integrity assessment for external corrosion using high resolution MFL reported a significant population of metal loss features (see Table 1).

Responding to the ILI predictions based on ASME B31G would not be considered for this case due to the impracticality for resourcing prompt response for a rehabilitation effort of this magnitude. Assessment of ILI and response using the B31G Modified criterion improves the situation but the response effort for 393 excavations could approach $10MM assuming a per dig budget of $25,000 USD with the potential for higher total costs depending on the variability of excavation costs. Pipeline operator experience with application of the B31G Modified criterion is that 30% to 50% of features identified by ILI may not require repair upon direct examination. Therefore given the economic risk a third option was evaluated.

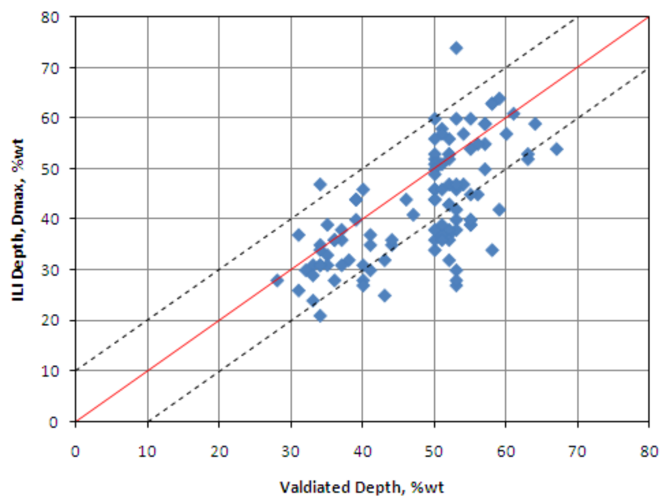

Figure 4: Depth Unity Plot for the Case Study

Assessment of ILI results and response using EAA criterion reduced the predicted repair locations to 63 excavation sites based on safety factor only, depth criterion was considered separately. Considering the sensitivity of EAA-ILI predictions to ILI measurement performance, rigorous performance evaluation was implemented in order to insure the measurement performance was sufficient to insure the prescribed factors of safety.

The in-ditch validation protocol included provisions for locating and matching of ILI features. The results of the depth validation are shown in Figure 4. The results indicate a small non-conservative bias in depth predictions; however, the depth performance claimed by the ILI vendor (80% certainty of +/- 10%wt) cannot be rejected based on the data. An apparent depth performance of 80% certainty +/-15%wt can be accepted with 95% confidence based on Clopper-Pearson confidence interval with true tool performance likely better when the in-ditch error is taken into account. [7] Comparisons were also made between the failure pressures predicted from ILI and the failure pressures measured in the field with the results shown in Figure 5.

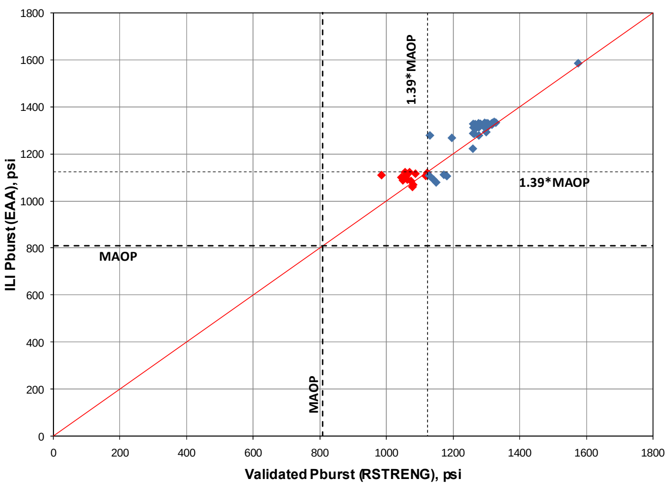

Figure 5: Unity Plot for Pburst, EAA failure predictions compared with RSTRENG from direct examination.

The results shown in Figure 5 are typical for high resolution MFL using Effective Area Assessment when claimed depth sizing performance of 80% +/- 10%wt cannot be rejected. The average for the ratios of predicted failure pressure to RSTRENG failure pressure was 1.02 with a standard deviation of 0.043. The mean value for the error (ILI-Field) exhibits a slight non-conservative bias (29 psi), with low standard deviation for the errors (28 psi). Application of a confidence interval to the differences in ILI and Field burst pressure predictions accepts +/-200 psi burst pressure variation at 95% confidence and 95% certainty. [7] Assuming the validation population is representative of the remaining metal loss feature population, ILI features accepted at the prescribed safety factor (1.39 for this case) have a low likelihood of exhibiting a safety factor less than one. The availability of rigorous validation data also offers the potential for calibrating the ILI data to provide more conservatism when needed to support risk targets.

In this manner, a higher accuracy ILI assessment criterion combined with consideration of ILI performance validation was utilized to determine a safe and effective integrity response for this pipeline.

Integrity assessments predicting large, critical metal loss defect populations are becoming less common as Integrity Management Programs mature. However, significant numbers of ILI re-assessments continue to be executed annually to assure future integrity. While such re-inspections should necessarily be less sensitive to ILI performance, verification of performance should still be considered as an acceptable alternative to full validation.

Case Study 2

API 1163 (Section 9) describes the methods that can be applied to verify that reported inspection results are within the ILI performance specification. Within the Standard a distinction is made between results verification with and without field verification measurements. It has been common within the Industry to perform very limited field measurement validations and statistically such efforts are meaningless. However, within the context of API 1163, one possible approach to address large scale tests (opening the way to verify ILI results without validation excavations) was investigated by one of the co-authors. API 1163 offers that reported results can be considered verified by comparison with the results from prior validated inspections on other lines, provided (1) the prior data represents the range of reported anomaly types and characteristics, and (2) the prior essential variables match those used in the current inspection.

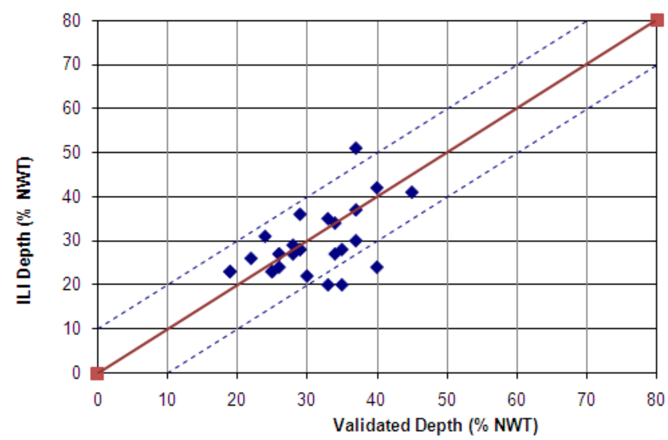

Figure 6: Unity Plot for validation data pooled from two similar segments inspected with the same inspection system

For transmission gas pipelines the principal threat is often proven to be external corrosion. The large scale of such systems often means that multiple valve sections are characterized by the same coating systems, pipe grade and year of construction offering an opportunity to consider that validation results from obtained from one section could be applicable to other segments inspected by the same ILI system. Figure 6 is a depth unity plot for validations from two separate ILI assessments of similar pipeline segments (service, coating, age, grade) using the same inspection system. A total of 27 validation comparisons were pooled and the vendor's specified depth sizing performance (80% +/- 10%wt) were within specified limits.

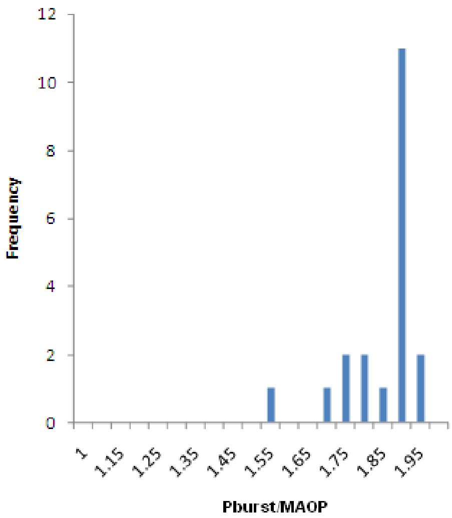

Figure 7: Distribution of safety factor predicted by ILI for an assessment predicting very few metal loss features.

A subsequent ILI assessment was conducted for a pipeline segment with characteristics similar to those for which pooled validation data was available. There were insufficient features greater than 20%wt (4 corrosion features) to satisfy the conditions necessary to validate to 95% confidence a certainty of p=0.8. Increasing the validation population to include features less than 20%wt deep would increase validation error due to difficulties to reliable location and matching of small features of severity similar to line pipe mill features. This subsequent ILI assessment reported no features that constituted an integrity condition. Figure 7 show the most severe feature exhibited a predicted safety factor of 1.55. Metal loss features accepted at a safety factor of 1.55 have an extremely low likelihood for actual safety factor less than 1 due to measurement error consistent with the stated ILI performance. Applying the principals of API 1163 - specifically the verification process, the ILI system performance based on the pooled data, and the sensitivity of the pipeline to integrity conditions and ILI performance, a case can be made that validation excavations are not required.

As integrity management programs progress, with improved mitigation measures employed to reduce total risk, it should be expected that ILI reinspections should report fewer and less severe metal loss features. The value of leveraging rigorous benchmark performance studies, pooling validation data and employing the verification process proposed in API 1163 has the potential for significant cost savings for programs with significant numbers of annual ILI assessments while maintaining consideration for safety.

SUMMARY

The goals of integrity management programs are to reduce risk posed by pipelines to the public and environment as well as reliably deliver products to their end customers. Operators must necessarily consider how to achieve these goals in a cost effective and sustainable manner and simultaneously provide the necessary assurances to stakeholders. When integrity management resources are overly focused on a few “significant” threats, the risk of other threats can be easily overlooked. The industry is becoming equipped with a more complete range of tools and techniques to be able to respond to ILI assessments in a safe and cost effective manner. Following the principals of API 1163, verification of performance for measurement systems is an important consideration in order to quantify the likelihoods of false acceptance of metal loss features as safe.

Metal loss sizing performance consistent with +/- 10 % wall thickness at 80% certainty can support the most widely used condition assessment criteria (ASME B31G, 0.85dL and Effective Area Assessment). The effectiveness of the assessment criteria in terms of safety and the management of response to ILI indications can be affected by actual ILI sizing performance. Full consistency with the tightest claimed specifications is not necessarily a requirement in order to manage safety. The key is to understand the sensitivity of the pipeline condition to measurement error in order to determine the level of effort required to validate ILI tool performance.

REFERENCES

1. ASME B31.8S, “Managing System Integrity of Gas Pipelines”, section 7- 2004.
2. Kiefner, J.F., Vieth, P.H., “A Modified Criterion for Evaluating the Remaining Strength of Corroded Pipe”, PR 3-805, Battelle, 1989
3. McNealy, R, Gao, M., Limon, S, Deaton, B., “Defect Assessment Using Effective Area Method from In-Line Inspection Data”, IPC2008-64481, Calgary
4. API 1163,” In-line Inspection Systems Qualification Standard”, API 2005
5. Haines, H., Johnston, D., “Using Effective Area Method with ILI Data Can Reduce the Uncertainty in the Resulting Calculated Failure Pressure”, Corrosion 2010, No. 14990, San Antonio, 2010.
6. Hallen, J., Caleyo, F., Alfonso, J., et al, “Statistical Calibration of Pipeline In-Line Inspection Data”, ESIQIE-IPN, PEMEX E&P, 16th WCNDT, Montreal, 2004
7. McNealy, R. McCann et al, “ILI Validation III Effect of In-Ditch Errors in Determining ILI Performance”, IPC2010-31269, Calgary, 2010
8. Tomar, M., Fingerhut, M., Yu, Deli, “Qualification of ILI Performance in Accordance with API 1163 and the Potential Impact for Management of Pipeline Integrity”, IPC2008-64469, Calgary, 2008