Increasing Confidence in Direct Assessment with Data Analytics and Large-Standoff Magnetometry

Lewis Barton, Mark Wright
ROSEN Group

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

Organized by
CLARION
TECHNICAL CONFERENCES

https://doi.org/10.52202/078572-0079

https://doi.org/10.52202/078572-0079

Abstract

In-Line Inspection (ILI) is widely used to monitor pipeline conditions, identify and size defects, and meet regulatory requirements. However, alternative approaches for 'non-piggable' pipelines are needed where ILI is not feasible. In these cases, knowledge-based models relying on data, engineering assessments, and assumptions are required.

External Corrosion Direct Assessment (ECDA) is a process where variables believed to contribute to corrosion are combined with above-ground surveys or computational modelling to identify corrosion 'hotspots' for in-field investigation. However, these techniques have several known limitations, often leading to wasted excavation campaigns and lingering uncertainty.

To enhance the direct assessment process, ROSEN has incorporated predictive analytics from its Integrity Data Warehouse (IDW). The IDW contains data from over 26,000 in-line inspections, covering more than 620,000 miles (1,000,000 kilometres) globally. These data provide significant improvement in predictive capacity and the likely condition of pipeline assets across all diameters, pressures, and fluids.

ROSEN has also begun incorporating Large Stand-off Magnetometry (LSM) into its direct assessment approaches. LSM detects changes in the magnetic signature of the pipeline that correlate with increased stresses, enabling the detection of a wide range of stress-raising anomalies and defects. This adds another layer of information, allowing for greater confidence in the identification of potential excavation sites.

This paper provides a detailed overview of the new process and the improvements made through the addition of LSM and the integration of the IDW into a ECDA process.

 https://doi.org/10.52202/078572-0079

Introduction

The process of External Corrosion Direct Assessment (ECDA) has been used throughout the pipeline industry, long before the introduction of the NACE standard SP502 in 2010. Primarily aimed at unpiggable pipelines, ECDA has been regarded as an industry accepted approach for assessing the corrosion condition of pipelines, where pigging is not possible. Although ECDA does have many advantages, in that it is a pre-emptive approach, it is heavily reliant upon the quality and quantity of data available. In order to be considered effective and robust enough to provide confidence in any results, users of the standard can be required to collect large volumes of costly data – ultimately, it is the volume of data collection, which is seen by many as the limiting factor of the approach. In addition, data collection is focused primarily on corrosion detection using Cathodic Protection (CP) measurements, which can be effected by ground or coating condition. Furthermore, although corrosion is a key failure mode, it only accounts for 28% of failures [1] leaving threats like ground movement, or physical, third party damage unaccounted for.

Data collected by In-Line Inspection (ILI) over many years contains information on corrosion trends across thousands of pipelines around the world. This data can be used to make predictions, regarding the likely condition of pipelines that have not been inspected. Historically, integrity and corrosion engineers have developed knowledge of the fundamentals of corrosion by academic study and learnt from experience with different pipelines.

Today, machine-learning models can be trained to make predictions of pipeline condition. The training of these models is typically most effective, when large volumes of good quality relevant data are available. Historic ILI data, particularly when combined with data such as rainfall, soil type and coating, provides a basis for predictive model development. Over the past few years, ROSEN have been exploring the machine-learning possibility, by collecting relevant datasets combined with our many years of ILI data history and GIS expertise. The result is an approach that delivers different types of models, depending on the application and available data. Generically we refer to these models as 'Virtual-ILI' (V-ILI), with the intention being able to make a prediction of what we would find by an actual ILI, if it were completed. The models range from low resolution basic predictions of overall condition for a pipeline, based on simple inputs such as construction year, coating, diameter, country, etc. Up to complex models that will predict the condition of specific segments, using multiple additional inputs including rainfall, soil types, environment, land-use, etc.

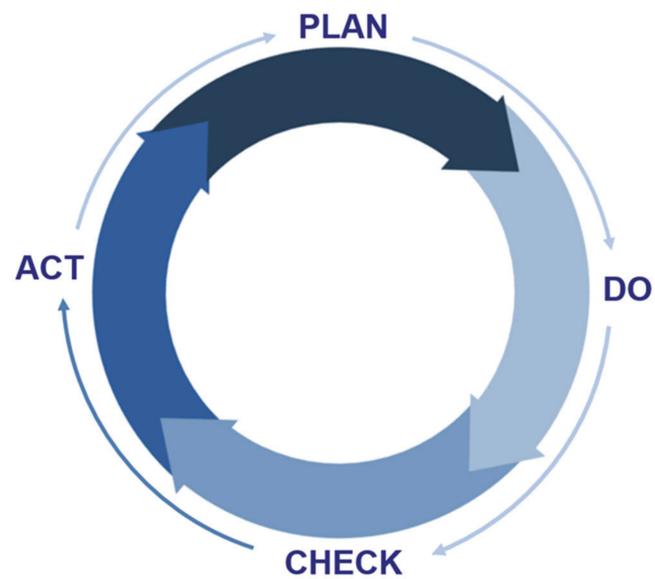
This paper discusses the use of the ROSEN V-ILI tool in order to enhance the ECDA process in combination with LSM data. A case study is also presented, where V-ILI and LSM was used as part of the ECDA assessment of a pipeline to provide not only further input data to the process; but also provide higher confidence – reducing the number of excavations required. This combination of inspection techniques is referred to as Non-Intrusive Pipeline Assessment (NIPA) within ROSEN.

https://doi.org/10.52202/078572-0079

The ECDA Process

The process of ECDA is founded on the simple integrity management loop of Plan, Do, Check, Act, and consists of four stages. A high-level summary of the four stages has been provided below:

- **Stage 1 – Plan – Preassessment**
    - Data collection and initial analysis, to decide on the inspection methods to be used and most importantly, if the ECDA approach is feasible.
- **Stage 2 – Do – Indirect Inspection**
    - Perform indirect inspection of the pipeline by desktop study and above ground surveys, to understand the likely condition of the pipeline. In order to identify and rank areas where external corrosion may be present, known as hotspots.



- **Stage 3 – Check – Direct Examination**
    - Excavation at the hotspot areas to confirm or deny the presence of corrosion, classifying and sizing corrosion anomalies, as well as the conditions to which the pipeline is exposed.
- **Stage 4 – Act – Post Assessment**
    - Review of the results from the ECDA process in order to assess the overall effectiveness of the analysis and perform a fitness for purpose / service assessment. Culminating in the definition of the reassessment interval and recommending integrity and corrosion management actions.

Although ECDA is an industry recognised approach, it does have limitations, namely it can only be used to detect corrosion and only where an indication may be reliably identified using CP detection methods such as Close Interval Potential Survey (CIPS). This means that the ECDA process can inadvertently lead to a false indication of the pipeline condition, especially if CP shielding is also present.

Through the addition of LSM to the ECDA process, further defect types can be identified, measured and the effects understood. Using a process of overlaying data sets to look for commonality between completely different phenomena enables higher confidence in results and a reduction in the chance of false negatives. Likewise, by further incorporating V-ILI data comparisons using patterns and

https://doi.org/10.52202/078572-0079

learnings from thousands of pipelines subjected to ILI, more can be done to compare and contrast the uninspected pipeline – Giving a prediction of condition based on the behaviour of similar lines.

Large Standoff Magnetometry

Large Standoff Magnetometry (LSM) is a non-intrusive inspection technology used for aboveground, indirect pipeline inspections. The sensor technology uses the phenomenon of magnetostriction, where a ferromagnetic material changes its shape or dimensions when subjected to a magnetic field. This deformation occurs because the magnetic field alters the alignment of magnetic domains within the material, causing it to expand or contract. The magnetic field in the case of pipelines can be that induced naturally from the earth and pipeline operation (passive), or applied artificially using AC signal generators (active). Similarly, the Villari effect, or inverse magnetostriction, describes changes in a material's magnetic susceptibility when mechanical stress is applied.

Using LSM identifies what is known as Stress Concentration Zones (SCZ), without having to physically change or intrude into the system, making it particularly useful for pipelines that are not possible to inspect using conventional methods, like ILI. LSM measures geomagnetic flux around steel pipelines in X, Y and Z vectors, across a sensor array, to detect and evaluate changes in stress states and indicate anomalies.

The presence and sizing of SCZ in a pipeline can be an indication of anomalies attributed to deformation, buckling, corrosion and ground movement, essentially anything that could change the uniform stress in the material. Claims have been made that LSM can identify phenomena such as Stress Corrosion Cracking (SCC); however, there is still no conclusive proof that LSM is a reliable detection method and remains to be confirmed. Nevertheless, LSM can be used to identify areas at higher susceptibility of SCC, moreover, it has been employed in ground movement monitoring to observe and check change in the shape of pipeline routes, perform bending strain analysis and the development of SCZ in vulnerable areas.

Although not specific to direct pipeline anomalies, LSM can also be used to detect and map underground pipeline furniture, such as valves, drains, foreign pipelines or cables and even stuck pigs by detecting changes in the surrounding magnetic field caused by increases in magnetic mass or stress.

Although LSM has shown its effectiveness in maintaining pipeline integrity, there are still many debates about its suitability as a standalone inspection method. This has not been helped by the lack of regulation in the area of LSM; the claims which can be made about detection capabilities, or the intellectual property rights guarding the analysis of data - meaning its results can appear “black box”. It is for this reason that ROSEN utilises LSM as part of a combination of inspection and analysis techniques, and not as a standalone solution, particularly in cases where high consequence areas are found. For example, when LSM is combined with direct assessment it serves as an enhancement

https://doi.org/10.52202/078572-0079

inspection, providing information on a lot more than simply corrosion. This is key to further refinement and ranking of anomaly indications, reducing unknown unknows and false positives or negatives – ultimately providing more confidence regarding pipeline integrity.

LSM comparison to ILI

ROSEN has performed a number of comparisons of LSM results to direct field verification, but also ILI, principally to validate performance and to further understand the limitations of the technology. The below Table 1 shows one of the comparisons made by ROSEN where the Estimated Repair Factor (ERF) was calculated for each defect detected in an ILI data set. These ERF values were then compared to the number of SCZ anomalies and their locations that were detected and measured by LSM. The ERF was chosen to be used as a measure of detection capabilities, as ultimately it is the ERF value that is used to determine the criticality for defect repair when ILI is performed.

The ERF is a numerical value used to prioritise and assess the need for repairs or mitigative actions on pipeline defects. It is a critical parameter in pipeline integrity management, particularly for evaluating the severity of metal loss or other anomalies detected during inspections and is typically calculated using industry-recognized defect assessment models, such as Modified ASME B31G and RSTRENG (Remaining Strength of Corroded Pipelines) [2]. These models consider factors like the defect depth, length, and shape, as well as pipeline operating conditions, expressed as the ratio of the pipeline's maximum allowable operating pressure (MAOP) or operating pressure to the predicted failure pressure of the defect.

The values calculated by the ERF are classified as

- ERF > 1.0: - The defect is critical and poses a risk of failure at the operating pressure. Immediate repair or mitigation is required.
- ERF ≤ 1.0: - The defect is not currently critical, but monitoring or future mitigation may still be necessary.

By systematically applying the ERF, pipeline operators can focus their resources on the most critical areas, minimising the risk of pipeline failures and ensuring long-term operational safety. Therefore, the comparisons to ILI were focused on identifying if LSM can reliably identify the most critical ERF defects. Providing this was the case, then there is confidence that although low level anomalies may be missed, significant anomalies which pose a threat to integrity would reliably be detected.

The findings of Table 1 show that this is the case, in that the Probability of Detection (POD) of smaller ERF defects is lower with LSM; however, larger defects (where the ERF is approaching or exceeding 1.0) were found in all cases. Although a slight increase in the number of defects detected was observed by increasing the distance tolerance, generally all defects were found at the same position as ILI. This provides high confidence that critical defects are reliably detected at the location they are active.

 https://doi.org/10.52202/078572-0079

Table 1: Comparison of ILI defects and LSM anomalies as a function of ERF

ERF	ILI Features	Corresponding LSM indications within distance tolerance bands							
		0 m tolerance		±1 m tolerance		±3 m tolerance		±5 m tolerance	
		Qty	%	Qty	%	Qty	%	Qty	%
<0,8	1172	492	42	530	45	624	53	669	57
0,8 to 0,9	13	6	46	7	54	8	62	8	62
0,9 to 1,0	1	1	100	1	100	1	100	1	100
1,0 to 1,1	1	1	100	1	100	1	100	1	100
>1,1	2	2	100	2	100	2	100	2	100

Virtual ILI

Virtual-ILI [3] is the process of using machine learning methods to learn from a global database of pipeline inspection information, for the purpose of predicting the likely condition of an unseen pipeline, one that has either yet to be inspected, or cannot be inspected with conventional ILI tools (Figure 1).

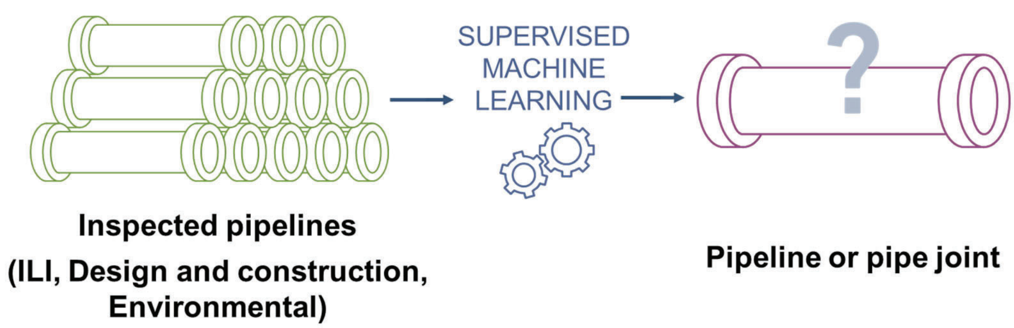


Figure 1: The Fundamentals of Virtual-ILI

A pipeline (or pipe joint) has a number of parameters that describe it, which include design, construction, and location information. These are used as predictor variables, and they form the basis of the inputs for the machine learning models.

Previous studies [3, 4, 5] have shown positive results using Virtual-ILI to predict third-party damage, and the density and maximum depth of external corrosion anomalies. In addition to this, generalised corrosion growth rate distributions that can also be applied to pipelines with similar location and construction attributes. Models have been trained to predict two external corrosion condition metrics; (i) maximum depth (% of wall thickness), and (ii) number of external corrosion defects per square meter.

https://doi.org/10.52202/078572-0079

Three variations of Virtual ILI are utilised to predict the condition of the target pipeline. These models are defined as:

- **Model A**

A basic model, trained on a limited number of predictor variables with the intention of giving a general overview of the pipeline condition based mainly on trends that relate to pipeline design and construction.

- **Model B**

A more sophisticated model, with environmental predictor variables in addition to the basic design and construction inputs. As with "Model A" the intention is to give a general overview of pipeline condition, but more accurate than using design and construction information alone.

- **Model C**

A further extension of "Model B", that segments the pipeline and delivers a per segment condition prediction. The predictor variables are the same as Model B, namely design, construction, and environmental data. The intention is to predict which segments are likely to be in better or worse condition, reflecting the reality that many pipelines are in generally good condition, and some have a few bad segments.

Dataset

For a model to be trained and evaluated, sufficient metal loss ILI inspection data representative of the target population that the Virtual ILI is attempting to predict must be available. For example, if we are trying to predict the condition of uninspected pipelines installed during a certain period, then it is important that the training data have enough of these groups to learn from. The same logic applies to other categories, such as external coating, pipe grade, location, etc. An imbalanced split of data between these groups (e.g. if the data is dominated by pipelines with a particular coating) can result in biases, with detrimental effects on the model's ability to successfully make predictions.

The Integrity Data Warehouse (IDW) is used to provide the data for this study. The IDW is a central repository containing in-line inspection data from tens of thousands of pipelines that ROSEN has inspected over multiple decades, including associated pipeline meta-data. Table 2, summarises the status of the IDW with respect to metal loss inspections at the time of writing.

Table 2: Integrity Data Warehouse Summary

Inspection runs	24,799
Number of pipelines	11,051
Number of pipe joints	66,604,244
Inspected length (km)	803,805

https://doi.org/10.52202/078572-0079

Number of external metal loss anomalies	91,497,322

Training data

Within the overall pipeline IDW, the dataset used for this study comprises of data from 1,868 pipelines; considered to be a subset with good representation for the predictor variables used as input. This includes pipelines from Europe and North America, with construction years ranging from 1940 to 2020.

The variables selected for model A are pipe joint design information including construction year, length, wall thickness, diameter, pipe grade and coating type. The model A is an entry-level model, providing a rough prediction of the overall condition of the pipeline based on a limited number of variables.

In order to improve model performance, additional geospatial and environmental meta-data is collected from external datasets and spatially joined with the location of each pipe joint. This process, known as geo-enrichment (Figure 2), is possible when the location of the pipeline right of way is accurately known. Both the pipeline geo-enriched (B) and segmented geo-enriched (C), Virtual-ILI models use this data in addition to the predictor variables used to train model A. Variables obtained in the geo-enrichment process include the following:

- Land use
- Terrain
- Soil type
- Local rainfall
- Intersections (water, road, rail, power line)

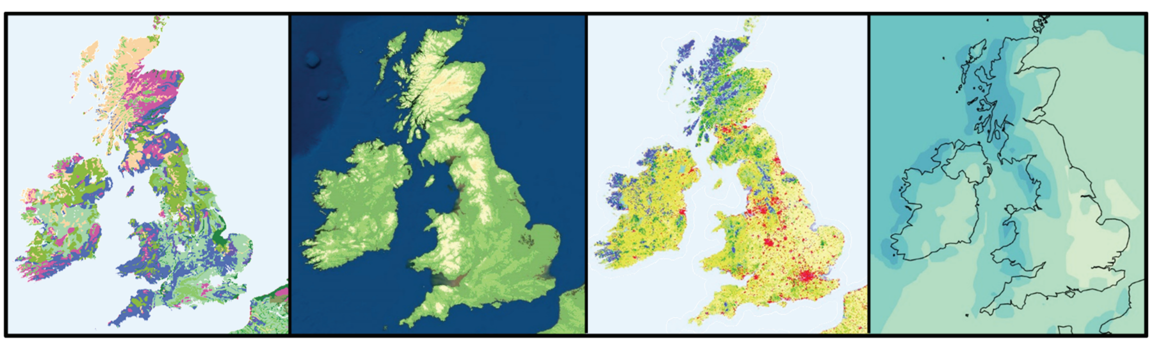


Figure 2: Geo-enrichment examples – from left to right: soil type, ground elevation, land use and average rainfall.

Feature engineering is used to provide additional variables which include location and climate, engineered features allows the model to learn which pipelines are located in similar regions of the world, both in terms of their actual location and the classification of the climate in which they reside. By engineering a variable, we are ensuring that the predictor variables are linked to one another.

https://doi.org/10.52202/078572-0079

The addition of these variables aims to increase both model accuracy and confidence of predictions by enabling the model to identify additional trends within the data, with respect to external threats giving rise to external corrosion. Similarly, having this data to pipeline joint resolution enables prediction to the level of specific pipeline segments or joints.

Methodology

The objective of supervised machine learning is to learn from data where the ground truth is already known, and then attempt to predict the probability of correctly identifying a specific class of ground truth on completely unseen data. For this study, the Virtual ILI models are trained to predict two condition metrics:

1. External corrosion max depth, an indicator of the severity of corrosion; and
2. Number of external corrosion defects per square meter, an indicator of the extent of corrosion.

Both target variables are calculated from ILI metal loss reports and then assigned into four categories (Table 3). The predictor variables relating to these inspections are then used in the model to find trends and underlying patterns that match patterns within these condition metrics. This is achieved by defining a function, f, that maps a set of predictor variables $\{x_i\}$, to a target variable, y, thus

$$y = f(x1, x2, \ldots, xn)$$

Table 3: Target variable categories

Parameter	Cat.	Value	Parameter	Cat.	Value
Number of defects per m²	1	0	**Maximum defect depth (% wt)**	1	0%
	2	>0 to 0.001		2	>0% to <25%
	3	>0.001 to <0.03		3	25% to <50%
	4	≥0.03		4	≥50%

Segmentation

Model C is designed to make per segment predictions, which are expected to be useful for integrity management purposes, particularly when considering external survey actions. Previous studies [4, 5] have made use of unsupervised learning techniques to segment the pipelines based on geospatial properties such as land use, intersections, population density, pipe grade. The use of unsupervised learning (hereafter referred to as clustering) to segment the pipelines allows an automated data-driven approach, that can deal with large volumes of data.

The case study below, uses an agglomerative clustering model (based on a hierarchical clustering algorithm) [6] to segment the pipelines. This takes input data for each pipe joint (including soil type and land use), and then determines clusters of closely related inputs, based on these predictor variables. The soil type and land use spatial feature variables yield 27 distinct categories, which we then use for classification purposes. Having this number of variables will result in long data processing times and difficulty with visualising and validating the results. To remedy this, an extra step is added to this process, which is dimensionality reduction, using a technique called Principal

 https://doi.org/10.52202/078572-0079

Component Analysis (PCA) [7]. PCA is used to reduce these variables to a low dimensional space, whilst also retaining a significant proportion of their explanatory power. Figure 3 illustrates an example of a pipeline being segmented, where S refers to the segment ID.

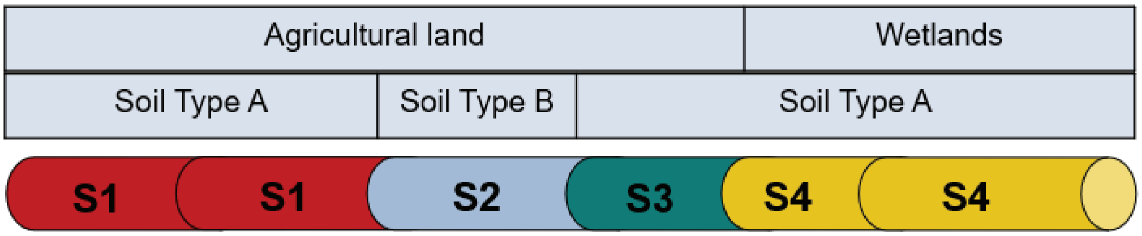


Figure 3: Example segmentation of pipeline

Model Selection and Validation

To demonstrate the Models for this case study, the parameters of the pipeline discussed below were used. The target variables are categories (ranges of values for maximum defect depth and defect density), which are referred to as a classification problem in machine learning. It is therefore necessary to select an algorithm that best suits the problem. Prior to any training, tuning or testing, 20% of the pipelines in the relevant dataset are set aside for the purpose of validation, also referred to as the serving set. This method ensures that the models are not tuned to this data and therefore will best represent how the model would perform, on unseen data.

Five machine learning algorithms were tested as part of the initial step, these were: Random Forest, Support Vector Machines (SVM), LASSO regression, MLP Classifier, and Extreme Gradient Boosting (XGBoost). 5-fold cross validation [8] was applied to these, where data is randomly partitioned into 80% training and 20% testing, where the test subset changes for five iterations until all the data has been used for training and testing. It is important to mention that within the cross-validation approach for the segmented resolution, we ensure that segments belonging to the same pipeline are never in both training and test set. In the author's experience, failing to do this can give a false increase in model performance due to the tendency for areas that are close together having similar values. This is known as Spatial Autocorrelation [9].

The most favourable results were achieved by the ensemble models (Random Forest and XGBoost), both knowing to perform well on complex datasets and be relatively robust to outliers within input data.

The next step in the selection process involves tuning multiple combinations of hyperparameters (~200) on both ensemble models. Hyperparameters are used to tune a learning algorithm with the goal of maximizing the model's performance in a given context. Examples of these include but are certainly not limited to, tree depth (how many questions will be asked in each tree before a prediction is made) or the number of trees/bootstraps to use. An important consideration when tuning a model is to ensure that it does not overfit.

https://doi.org/10.52202/078572-0079

Overfitting is the phenomenon of a model being tuned too closely to the training data, and providing poor results on test data. This typically results in a solution that does not generalise well to unseen data. Based on this, within the context of a random forest model, more trees or leaf notes are not necessarily the best choice, reinforcing the need to explore multiple hyperparameter combinations. Additionally, randomised thresholds can be used at each decision node in an ensemble of decision trees, which tends to enhance the overall performance of the ensemble by reducing the correlation between each specific model within that ensemble.

The performance metrics used to validate these models were accuracy, balanced accuracy (an average of each class accuracy), as well as the confidence level in the predictions being made. A summary of the best performing algorithms for each of the models is provided in Table 4.

Table 4: Selected ML algorithms for each model

Model	Target Variable	ML Algorithm
A	Number of defects per m²	Random Forest
	Max Depth (%)	XGBoost
B	Number of defects per m²	XGBoost
	Max Depth (%)	XGBoost
C	Number of defects per m²	XGBoost
	Max Depth (%)	Random Forest

The overall performance of the different models is summarised in Table 5.

Table 5: Model Performance

Model		Input Variables	Target Variable	Accuracy (%)	Balanced Accuracy (%)	Accuracy Within 1 class (%)
A	Basic	10	No. defects per m²	55	55	85
			Max Depth (%)	51	51	89
B	Geoenriched	22	No. defects per m²	63	62	92
			Max Depth (%)	56	57	92
C	(geoenriched + segmented)	22	No. defects per m²	63	57	90
			Max Depth (%)	63	51	94

"Accuracy" refers to the overall accuracy of correct predictions against incorrect predictions.

"Balanced accuracy" (Bal. Acc.) is calculated by taking the average of each class accuracy (a robust metric when there are class imbalances).

The last column denotes the confidence in which the model is making predictions on the test line based on what it has learnt from the training data. A high confidence would suggest that the model has seen pipelines sharing similar properties with consistent levels of corrosion.

Table 6 shows a confusion matrix illustrating predictions made for defects per m² using Model C where correct predictions refer to the numbers along the diagonal (predicted class and true class are

https://doi.org/10.52202/078572-0079

the same). The least favorable of these results are the top right and bottom left corner where the former refers to model predictions of high corrosion density when the true value is in fact no corrosion. Similarly, the bottom left shows model predictions of no corrosion when the true value is high corrosion. Each of these extremes collectively represent 3.1%.

Table 6: Confusion matrix showing performance of Model C in predicting feature density.

Accuracy 62.84%					
True Label	**No Corrosion**	183	37	20	9
	Low	21	35	22	1
	Mid	16	14	66	18
	High	7	0	26	39
		No Corrosion	**Low**	**Mid**	**High**
	Predicted Label				

Case Study

The combined NIPA and V-ILI approach was utilised for a pipeline that was consider to be typical for the application of ECDA. This case study present a summary of the findings. It is a relatively short pipeline, just 7km long that crosses agricultural land and is a relatively high pressure section of a gas distribution system taking natural gas from a national transmission system and delivering it to a small town. The pipeline was installed in the mid 1970's and during its operational life it had never been subjected to any inspection or pigging activities. A summary of the pipeline details is given in Table 7.

Table 7: Pipeline summary

Description	High Pressure Gas Distribution
Length	~7 km
Nominal Diameter	14 inch (377 mm)
Wall Thickness	9 mm
Pipe Grade	API B
Design Pressure	38 barg
Construction Commissioning Date	mid 1970's
Coating Type	Bitumen

Prior to this study, there was uncertainty regarding the condition of the pipeline. There had been no internal inspection, and the results of any historical above ground surveys had been mislaid. In addition to this, limited information with regards to the performance of the pipeline corrosion mitigation barriers, such as cathodic protection was available. However, there was no physical or direct evidence that the pipeline was in a poor or degraded condition. Instead due to the lack of confidence the risk attributed to the pipeline was driven by expectations of the operator, based on the pipeline age and lack of inspection confidence.

Typically to improve confidence in the condition of a natural gas pipeline, and demonstrate good practice to stake holders, options, as outlined in AMSE B31.8S [10], are hydrotest, internal inspection or Direct Assessment. However, hydrotest requires a pipeline to be taken out of service, which is disruptive and can be very costly. In addition, depending on the pipeline condition, the amount of associated repairs may become excessive and no information is gained regarding coating or pipe wall anomalies that could become a problem in the future.

In-line inspection provides detailed information on the pipe wall, allowing defects to be monitored and repaired if critical, or before they grow to unsafe dimensions. However, installing the equipment needed to launch and receive inspection tools, and controlling flow to achieve good inspection results can also be costly and disruptive. Direct Assessment is widely used but requires reliable input data (ideally historical records), excavations to confirm affected areas, and it is generally considered less reliable than internal inspection, but to improve overall confidence ROSEN used a NIPA approach with V-ILI.

The Stage 1 Pre-assessment concluded that, as the pipeline conveys dry sales gas for customer use, it was unlikely that internal corrosion was present; therefore, efforts should be focussed on an ECDA approach. A gap analysis showed that there was insufficient data to immediately move to Stage 3 and select locations to excavate and prove condition. The combination of the age of the pipeline (>40 years), and lack of reliable records, gave the expectation that the condition of the pipeline may be degraded or poor. Experience suggests that diligent operators who maintain their pipelines in good condition also keep comprehensive records, so expert opinion is inclined to caution when records are missing.

The next step, as part of Stage 2, was performing a LSM, Close Interval Protection Survey (CIPS) and Direct Current Voltage Gradient (DCVG) survey to gather information regarding the performance of the Cathodic protection polarisation, coating condition, environmental parameters and stress anomalies. While above ground surveys are generally easier to complete than ILI or hydrotest, it is not a trivial undertaking and achieving high quality results requires the mobilisation of an experienced team and access to walk the pipeline route, which can also be difficult to arrange and be costly.

Upon review of the above ground data, and corelation of anomalies, the results showed that while there were a number of possible defects found by the DCVG survey, the pipeline was fully protected by the ICCP system. However, in the absence of historical CP data, it was not known if the CP system had been providing sufficient protection over the past 40 years. Further comparison of the pipeline CP defects to that of LSM, found that there were no significant SCZ features in the immediate area of coating defects. Although there were some low level stress indications, they were not at an intensity that would indicate significant deterioration of the pipeline integrity. The results can be seen below:

 https://doi.org/10.52202/078572-0079

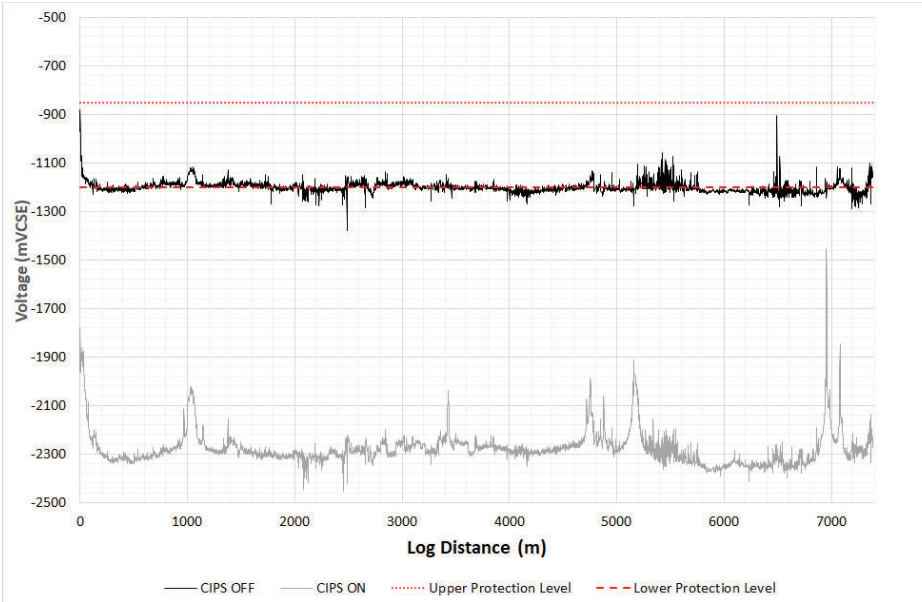


Figure 4: Close Interval Potential Survey (CIPS) results

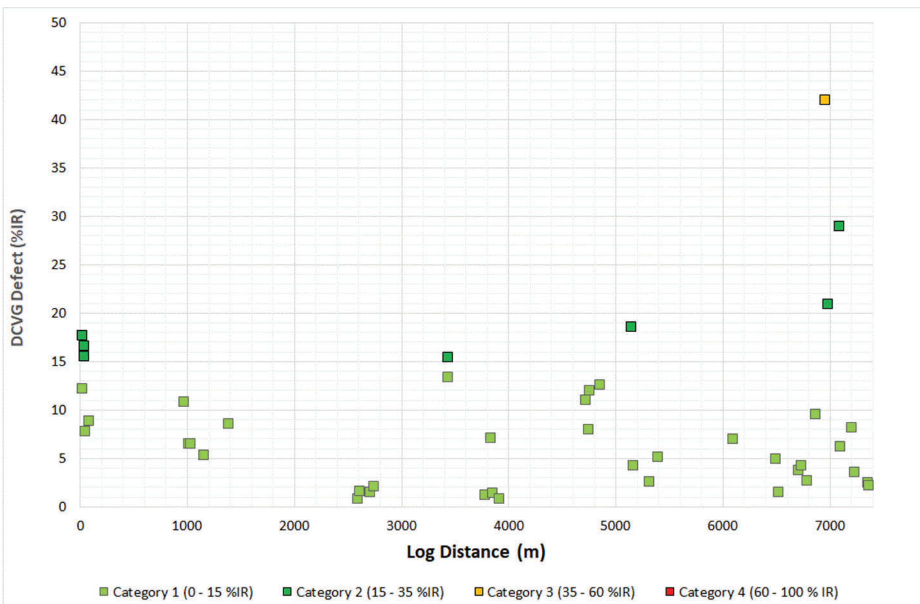


Figure 5: Direct Current Voltage Gradient survey results

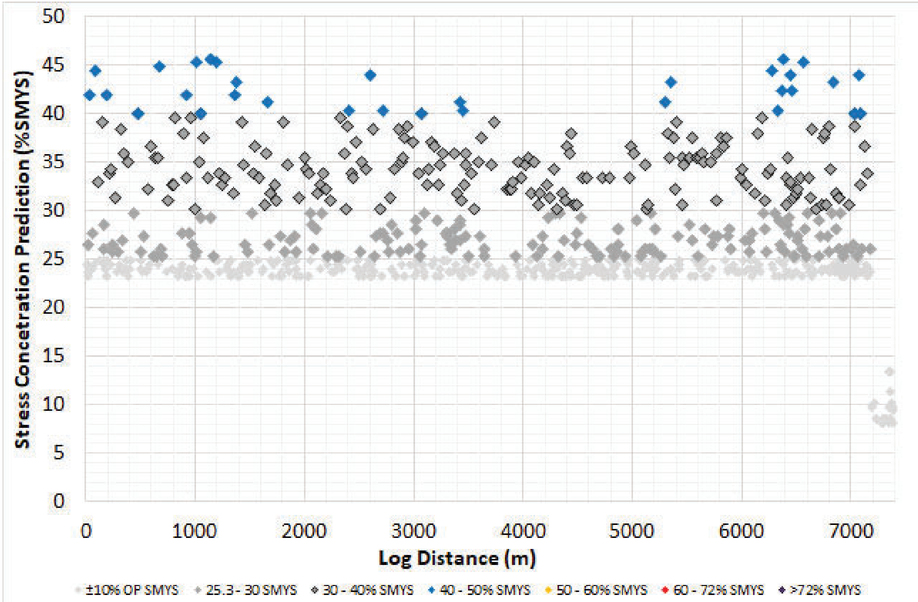


Figure 6: Large Standoff Magnetometry – location of SCZ defects vs. %SMYS

When these initial results were presented to the client, there was some doubt as to the reliability of the inspection. This was due to the stark contrast of the above ground inspection results and assessment, compared to the condition expectations of the client – driven by the age of the pipeline. In these situations, it can be the expectation of finding many defects that drives what is perceived to the be a successful inspection. Therefore, if an initial expectation is that a pipeline should have many features and none are found, it can be interpreted that the inspection must be incorrect.

To support the findings of the NIPA ECDA above ground work and further justify the result of the above ground surveys, Virtual-ILI was deployed.

Model A was run first to give an initial indication of likely condition of the pipeline. This lightweight model using only 10 design and construction predictor variables predicted the target pipeline defect density (defects per m2) to be >=0.001 - <=0.03 with a confidence of 53%. Maximum defect depth was also predicted to be between 25% and 50% depth metal loss, also with a confidence of 53%. To give some context this condition prediction is compared with the condition of all pipelines in the IDW in Figure 7.

https://doi.org/10.52202/078572-0079

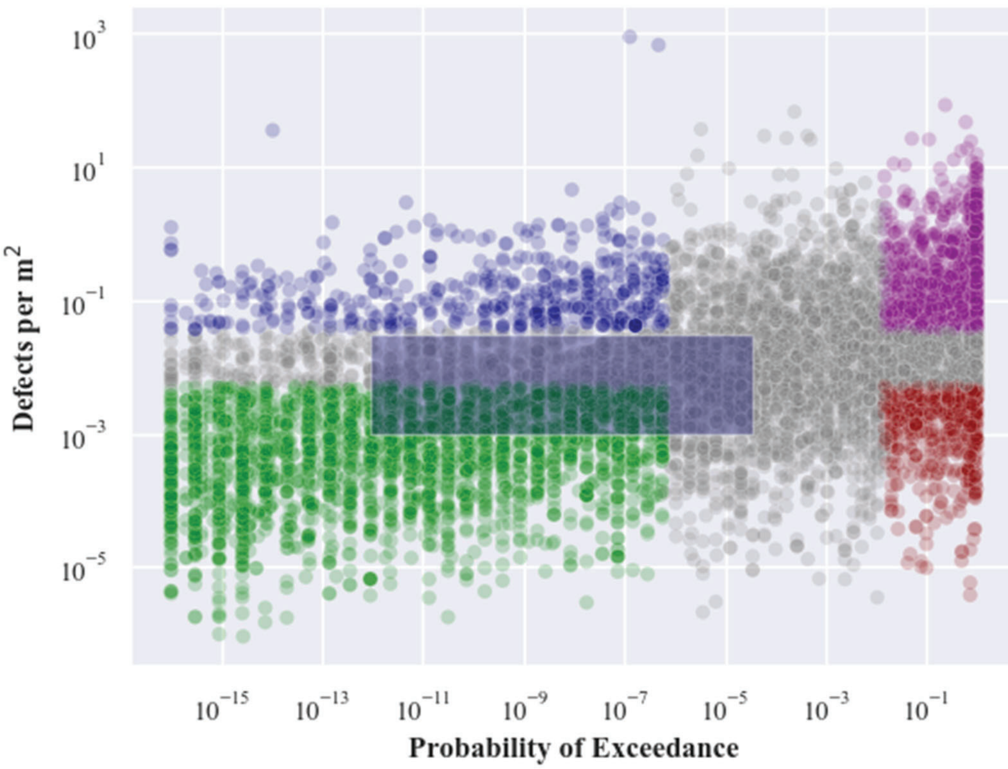


Figure 7: Model A (basic) Predicted condition class compared with full population of IDW

The condition as predicted by Model A covers a range of the total world population of pipelines. Note that the areas coloured green, blue, red and purple represent lower or upper quartiles with respect to probability of exceedance (probability that defect depth is > 80% wall thickness) and defects per m^2. The subject pipeline is predicted to be somewhere in the lower part of the population with respect to defect depth and around the middle of the population for defect density. This suggests that if the condition is similar to that of pipelines of similar, age, coating, grade, diameter, etc. then corrosion can be expected to be relatively extensive, but the depth should not be too severe. The confidence in the predictions was considered to be reasonable, given the limited input data used, but on the low side at 53%. Therefore, a digital model of the pipeline route was created, relevant environmental data was collected, and Model B was run.

Model B also predicted defect density to be >=0.001 - <=0.03 defects per m^2 but with a much improved confidence of 93%. Maximum defect depth was predicted to be between 0% and 25% of wall thickness, with a confidence of 61%. The Model B condition prediction is compared with the condition of all pipelines in the IDW in Figure 8.

https://doi.org/10.52202/078572-0079

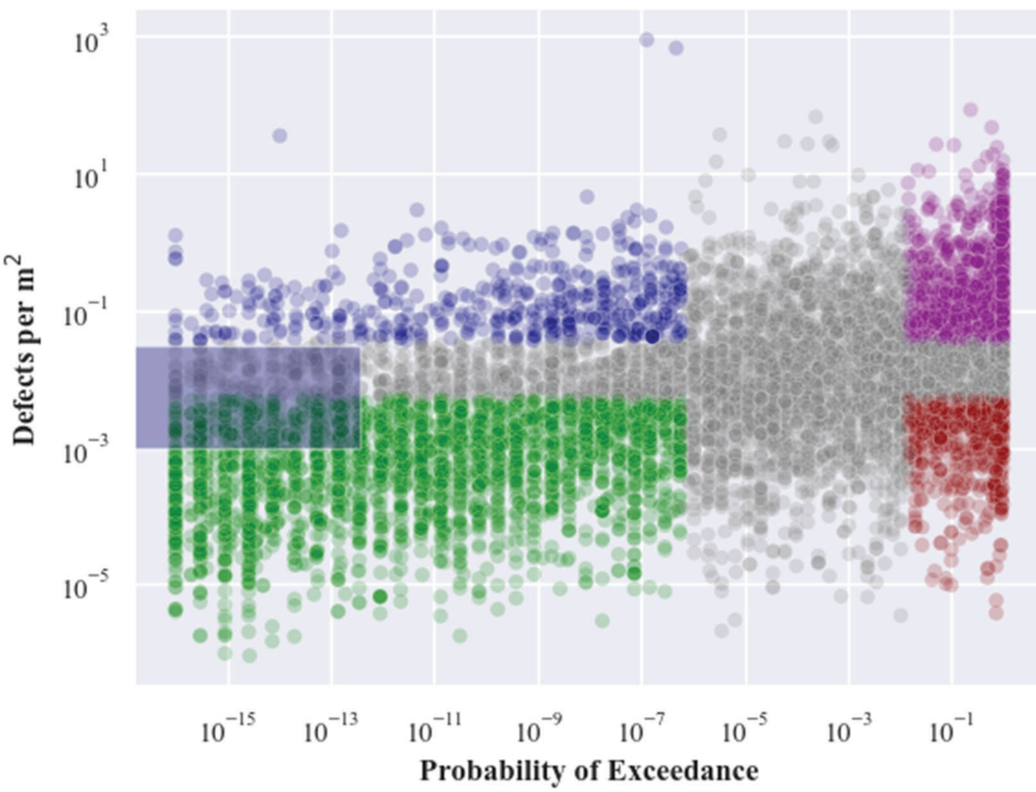


Figure 8: Model B (geo-enriched) prediction of condition class against full IDW

The condition predicted by model B is clearly better than that predicted by the basic Model A, with a lower class for maximum defect depth. The potential extent of defects however remains relatively high.

These results indicate that, assuming the subject pipeline condition is represented by similar pipelines in the training set, then an ILI inspection might find a relatively high number of defects, but it is unlikely that any would be very deep or need repair. This supports the finding of the above ground surveys, and confirms that the likely defects would be relatively minor.

The outcome of Stage 2 of the ECDA process is to define locations for excavation and direct examination of the pipe. As the V-ILI provided further justification that the number and severity of defects should be low, ROSEN used its NIPA process to define and rank locations for direct examination which would be representative of areas most likely to contain the most severe anomalies. In this type of situation, the number of excavations required to prove the condition of the pipeline can be substantial. Especially proving a lack of corrosion defects, which is inherently more difficult than proving that corrosion is present. However with the addition of LSM and V-ILI justification, the number and location of excavations can be much more easily justified.

https://doi.org/10.52202/078572-0079

In order to optimise the number of excavations required and to gain further confidence in the extent and severity of the corrosion, V-ILI Model C was utilised. The aim of Model C was to further segment the pipeline to identify how many possible segments, and if particular segments, would be more likely to contain corrosion. Thereby enhancing the confidence of the results from the direct assessment methodology, and provide further justification of the pipeline condition.

Model C (geo-enriched and segmented), identified two segments; Segment 1 running from the start of the pipeline to approximately the 6 km point. Segment 2 comprising the remainder of the pipeline. The segments are shown below in Figure 9 on a map. Blue and red refer to segments 1 and 2.

Figure 9. Map showing pipeline route with Segment 1 in blue and Segment 2 in red, plus excavation locations.

Feature density was predicted to be >=0.001 - <=0.03 defects per m^2 with a confidence of 80% for both segments, suggesting a uniform distribution of features along the whole length, qualitatively matching the results of the DCVG and LSM survey (note the V-ILI is run independently from these inputs). Maximum feature depth was predicted to be 0% - 25% wall thickness for segment 1 with a confidence of 32%, and the same for segment 2 with a confidence of 39%. This suggests that any deeper defects are predicted to be found in Segment 2, the last 1 km of the pipeline. Noting that the confidence in this prediction is low. Reasons for the low confidence were not investigated but could include:

- The training data was not sufficiently representative,
- The condition of the representative segments in the training data may have been highly variable, or
- Similar segments in the training data may be close to the edge of the defined thresholds.

Following analysis of the above ground survey data and the Virtual-ILI results, four excavation locations were chosen to give the best chance of finding any significant corrosion and best represent the possible data spread.

The distances and criteria are summarised in Table 8.

Table 8: Excavations resulting from the Above Ground Survey and V-ILI – location ID, associated distances and criteria.

Location ID	Distance (m)	Segment	Comment
1	7,079	2	Highest combination of factors (CIPS, DCVG, LSM and V-ILI)
2	6,953	2	Highest DCVG and LSM Defect, plus in V-ILI
3	1,147	1	Defect with low level DCVG and no significant SZC in segment 1
4	6,384	2	Control site in V-ILI segment 2

The combination of Virtual-ILI and above ground surveys all suggested that it was unlikely that deep corrosion would be found at any location. Locations 1, 2 and 4 all in segment 2 were considered to have the highest likelihood of having significant corrosion. While location 3 was required in order to perform validation in Segment 1, where the likelihood of deep corrosion was considered to be lower.

Excavation Results

Location 1

Contrary to the expected design, the coating was found to be a single layer Polyethylene (PE) Tape, not bitumen, casting further doubt on the system records. A single coating defect was noted due to soil loading of the wrap, coupled with poor adhesion, as the coating peeled away easily and light surface corrosion was visible on the pipe surface. Likely to be a result of poor surface preparation during the wrapping process as mill scale was removed and impregnated within the adhesive. The CP system confirmed to be working, evidenced by the white hydroxide deposits beneath the coating. Crucially there was no evidence of corrosion of any significant depth at the location.

Figure 10: General Excavation Findings at Location 1

Location 2

The coating was again found to be a single layer Polyethylene (PE) Tape not bitumen. Again, minor coating defects were noted. A single coating defect was found due insufficient overlapping of the

 https://doi.org/10.52202/078572-0079

wrap at the 6 o'clock position, coupled with poor adhesion furthermore attributed to poor surface preparation during coating application. The CP system was confirmed to be working, evidenced by the white carbonate deposits beneath the coating. Crucially, there was once again no evidence of corrosion of any significant depth at the location.

Figure 11: General Excavation Findings at Location 2

Location 3

The coating was a rubberised wrap coating not bitumen or the PE tape seen at locations 1 and 2, casting further doubt on the system records. Minor coating defects were discovered. There was also evidence of poor adhesion, as the coating peeled away easily and light surface corrosion was visible on the pipe surface at the interloping areas. The CP system was working, however it was clear some shielding had been present leading to the formation of some minor corrosion pits < 1 mm deep (< 11% of wt).

Figure 12: General Excavation Findings at Location 3

Location 4

At the final location, the coating was confirmed to be the original 1970's bitumen. Given the relative age of the coating and initial appearances, it was found to be in good condition, with no coating defects present. The bitumen was found to be brittle and easily removed, however this is to be expected from bituminous coatings of this age. Following removal of a small section of coating to confirm the beneath condition, the CP system was found to be functioning correctly, with a thin carbonate layer present on the surface and no evidence of corrosion of any significant depth.

https://doi.org/10.52202/078572-0079

Figure 13: General Excavation Findings at Location 4

Table 9: Summary of Excavation Results

Location ID	Distance (m)	Comment	Corrosion Presence	Comment
1	7,079	Highest combination of factors	No significant features found	PE tape, some coating defects, no measurable corrosion
2	6,953	Highest DCVG Defect	No significant features found	PE tape, some coating defects, no measurable corrosion
3	1,147	Defect with low level DCVG	Minor corrosion defects found <1mm depth (<11%WT)	Rubberised Wrap, coating defect at area of poor overlapping, possible shielding effect as a few small corrosion pits were found
4	6,384	Control site	No significant features found	Bitumen, no coating defects, no measurable corrosion

In summary, there were no significant corrosion defects at any of the 4 locations excavated. Three of the excavation sites were in the segment of the pipeline that Virtual-ILI predicted to be in the worst condition, and two at the locations of the most significant areas derived from the from above ground surveys i.e. CIPS, LSM and DCVG.

Conclusions

Through the use of Data Analytics (V-ILI) and Large-Standoff Magnetometry, there can be a direct increase in assessment confidence, when following an ECDA program as part of a pipeline integrity management plan.

 https://doi.org/10.52202/078572-0079

The success of many integrity analysis are often judged by how many defects are found; however, in some cases, success must be based on their absence – providing this matches the predictions made as part of the indirect assessment. Having confidence in a lack of features can be a challenge, but V-ILI allows expansion of the data horizon and instead view thousands of other pipelines. In addition LSM provides a further data set for analysis that isn't effected by the limitations of CP inspection.

Taking learnings from the case study , the pipeline (despite its age and lack of historical data records) was found in good condition and is fit for future service. Nevertheless, expert opinion alone would have concluded that the condition was uncertain and that potentially significant metal loss may be present. In the absence of historical records, the expert opinion was constrained and hence cautious due to the lack of relevant data. The V-ILI models developed using machine learning were based on a dataset of nearly 2,000 pipelines, which predicted the condition to be fair. That is, some corrosion was predicted, 0.001 to 0.03 features per m^2 (or up to 1 feature every 2.5 pipe joints), and maximum depths of 0 - 25% wall thickness. These Models were useful in supporting the ECDA process, most notably as part of the pre-assessment, with minimal initial data, through interpretation of above ground survey results and the selection and completion of relevant excavations.

In summary, The integration of V-ILI and LSM into an ECDA process provides data to back up the expertise and opinions of pipeline integrity / corrosion subject matter experts. Strengthening the position of the expert and providing them with an additional input, that can be used when historical inspection or survey data is sparse or questioned. This is especially true in the case of pipelines were minimal anomalies may be present, as proving the absence of defects can be more challenging than identifying their presence!

Further Work

The initial results of integrating V-ILI into the ECDA process as a screening tool show promise, especially to add further confidence in ECDA results when limited data is available. The integration of additional ILI data into the IDW, increases the variety and amount of relevant pipeline data from different and similar cases, also expanding the capability of the V-ILI for more accurate predictions. ROSEN will be further developing not only the model algorithms, but also how V-ILI can be integrated into the core of the ECDA process and complimented by LSM data.

Although useful in this instance, the primary use case seen for V-ILI is considered to be for preliminary pipeline screening, where operators have many pipelines that may need inspection – but only a finite budget. Subjecting a group of pipelines to V-ILI as an initial assessment, would allow for ranking of the most critical pipelines to be subjected to further inspection, with minimal input data.

https://doi.org/10.52202/078572-0079

References

[1] US DOT Pipeline and Hazardous Materials Safety Administration PHMSA Pipeline Incidents: (2004-2023)

[2] Pipeline Operators Forum, Specifications and Requirements for Intelligent Pig Inspection of Pipelines, version 4.3, 2022

[3] Capewell, M. Linsley, A. Martin, J. Taylor, K. and Palmer-Jones, R. (2022). *Virtual ILI – predicting the results of an In-Line Inspection.* Pipeline Pigging and Integrity Management (PPIM) Conference, February 2022, Houston Texas, United States of America.

[4] Carrell, S. White, J, Taylor, K. Martin, J. Irvine, S. and Palmer-Jones, R. (2022). *Predicting External Corrosion Growth Rate Distributions for Onshore Pipelines for Input Into Probabilistic Integrity Assessments.* The Digital Pipeline Solutions Forum (DPSF), May 2022, Houston Texas.

[5] White, J. Taylor, K. Martin, J. Carrell, S. and Palmer-Jones, R. (2022). *Estimating Pipeline Probability of Failure Due to External Interference Damage Using Machine Learning Algorithms Trained on In-Line Inspection Data.* International Pipeline Conference, September 2022, Calgary, Canada.

[6] Murtagh, F. and Contreras, P. 2012. *Algorithms for hierarchical clustering: an overview.* Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2(1), pp.86-97

[7] Abdi, H. and Williams, L.J., 2010. *Principal component analysis*. Wiley interdisciplinary reviews: computational statistics, *2*(4), pp.433-459.

[8] Fushiki, T., 2011. *Estimation of prediction error by using K-fold cross-validation*. Statistics and Computing, 21, pp.137-146.

[9] Legendre, P., 1993. *Spatial autocorrelation: trouble or new paradigm?* Ecology, 74(6), pp.1659-1673.

[10] ASME B31.8S, 2022 - Managing System Integrity of Gas Pipelines.

https://doi.org/10.52202/078572-0079