MAXIMISING ACCURACY OF MFL PIPELINE INSPECTION

Primary Author: Ben Scott, Baker Hughes, UK

Co-Author: Stephen Farnie, Baker Hughes, UK

Abstract:

There have been significant advances in magnetic flux leakage (MFL) in-line inspection (ILI) technologies in recent years. These have led to improvements in Probability of Detection (POD), Probability of Identification (POI) and Probability of Sizing (POS).

Whilst often the main focus of these advancements is the inspection vehicle itself, the end product of an inline inspection service is reliable and accurate data. This end product is influenced by various technological factors which include: recognition & detection algorithms; complex sizing models; robust and rigorous processes and; highly trained and skilled data analysts.

This paper explores all the main factors that contribute to delivering the reliable and accurate inspection reports that pipeline operators demand today. This review will be supported by extensive comparison of 'as reported' data vs 'in ditch' findings. This is particularly valuable for operators of offshore pipelines, where proving ILI performance is at least challenging, and often not possible.

Introduction

Pipeline operators must balance key concerns while running a business (Figure 1). Protecting people, the environment, and the reputation of the industry remain the highest priority, whilst maximizing the ongoing returns to shareholders from major investments is always a focus.

Figure 1: Key Drivers for pipeline owner/operator

These concerns become an even greater challenge in tough economic times with budgets continually under pressure. Many operators are choosing to collect a full, or enhanced inspection dataset but put priority focus and advanced analysis on targeted and problematic regions of interest. Such areas may be of high consequence or other regions identified from historic inspections or risk assessment.

Whether operators are focusing on targeted areas or conducting a detailed assessment on the entire pipeline section, the accuracy of the data used, in this case in-line inspection (ILI) data, has a significant impact on the outcomes of these assessment [ref 6].

Magnetic Flux Leakage (MFL) is the most widely used ILI technology in the world today. This is largely due to its ability to deliver in a wide range of pipeline operating environments whilst maintaining high levels of accuracy. Accuracy for MFL inspection is generally measured in terms of Probability of detection (POD), Probability of Identification (POI) and Probability of Sizing (POS).

The first commercial On-Line Inspection Centre (OLIC) MFL inspections took place in the 1980s (first British Gas MFL inspection took place in 1977). Over the four decades since, there have been significant advances in MFL inspection technologies and their resulting capabilities. Whilst often the focus of these advancements is the inspection vehicle itself, the end-product of an inline inspection service is **reliable and accurate data**. This end-product is influenced by various technological factors which include: recognition & detection algorithms; complex sizing models; robust and rigorous processes and; highly trained and skilled data analysts.

This paper will first highlight how MFL ILI 'accuracy' has changed and improved over time and then focus on the following factors which all contribute to the reliable and accurate inspection service. The factors covered will be:

	Accuracy impact
- The Inspection Vehicle	POD
- Software & Feature Recognition	POD & POI
- Data Analysts & Data Analysis Process	POI & POS
- Algorithms & Sizing Models	POS
- Performance Validation, Verification & Improvement	POD / POI / POS

Accuracy

MFL inspection accuracy is typically stated in terms of detection, identification and sizing. Each one of these is measured in terms of confidence levels, typically at 80 or 90%:

Detection or POD really means will 'it' be seen?
Commonly defined in industry by **API 1163 (American Pipeline Institute)** as *"The probability of a feature being detected by an ILI tool"* or by **POF (Pipeline Operators Forum)** as "*The probability that a feature with a size will be detected by the ILI tool."*

Identification or POI really means what is 'it'?
Commonly defined in industry by **API 1163** as *"The probability that the type of anomaly or other feature, once detected, will be correctly classified (e.g. as metal loss, dent, etc.)"* or by **POF** as "*The probability that a feature is correctly identified by the ILI tool."*

Sizing or POS really means what size is 'it'?
Commonly defined in industry by **API 1163** as *"The accuracy with which an anomaly dimension or characteristic is reported"* or by **POF** as "*Sizing accuracy is given by the interval with which a fixed percentage of features will be sized. This fixed percentage is stated as the certainty level."*

Detection and sizing specifications are typically a key element of an ILI contract. In some of the early inspection contracts from the 1980s, the accuracy levels were not stated or 'silent' largely because the specifications were unproven or did not even exist. Some reports would merely provide a distance and Asterix*, which effectively said 'there might be something here'. Although defects identified were effectively being reported on a reasonable endeavors basis, there was enough confidence in these results for MFL inspection to be of significant value in pipeline integrity management. This value contributed to improvements over the coming years and decades.

In the later 1980s and into the 1990s, detection and sizing specifications became the norm in ILI contracts. Figure 2 shows an example of a MagneScan contract specification. Specifications were provided for pits and general metal loss, with the minimum detection for pits @ 50% WT and for general metal loss @ 30% WT. This specification was commonly known at the time as '30/50 spec'. The sizing accuracy was +/- 20% or +/- 15% WT depending on the defect type.

ETECTION, SIZING AND LOCATION ACCURACY FOR 150MM TO 1400MM FOR SEAMLESS MANUFACTURED PIPELINES

	METAL LOSS CATEGORY		
	Pitting <(3tx3t)*	**General >(3tx3t)***	**Gouging**
Minimum Depth for Accurate Sizing	0.5t with surface dimension greater than: (t/2+10mm)x(t/2+10mm)	0.3t	If w>2t or 15mm**=0.5t If w>3t or 25mm**=0.3t
Sizing Accuracy *(Depth)*	±0.2t	±0.15t	If w>2t or 15mm**=±0.2t If w>3t or 25mm**=±0.15t
Sizing Accuracy *(Length)*	±10mm	±20mm	±20mm
Location Accuracy *(Axial)*	±0.2m between the feature and the reference girthweld and ±1% of stated distance between reference upstream girthweld and identification location reference		
Location Accuracy *(Circumferential)*	±7.5 degrees which for ease of reference is stated to the nearest half hour clock position		

Figure 2: Extract from a MagneScan ILI contract from the 1990s

In the 15-20 years that followed, the specifications improved and evolved to cover a greater range of defect, sizes, types (typically quoted according to POF feature category) and improved levels of accuracy. Table 1 below provides the inspection accuracy from the Baker Hughes fleet: This is the MagneScan fleet's, industry leading 'Super High Resolution Plus' (SHRP) specification. Today the minimum detection and sizing level is from 4% and +/-8% of local wall thickness, compared to 30% and +/-15% from 20 years earlier.

		General metal loss	Pitting	Axial grooving	Circumferential grooving	Pin hole	Axial slotting	Circumferential slotting
Reference dimensions (length x width)		4t x 4t	2t x 2t	4t x 2t	2t x 4t	0.5t x 0.5t	2t x 0.5t	0.5t x 2t
Super High Resolution *Plus*	Min. Depth At 90% POD	4%	6%	6%	4%	13%	13%	4%
	Depth Sizing accuracy	±8%	±8%	-13% +8%	-8% +13%	-13% +8%*	-18% +8%	-8% +13%
	Width Sizing accuracy	±12mm ±0.47 in	±12mm ±0.47 in	±12mm ±0.47 in	±12mm ±0.47 in	±7mm ±0.28 in	±12mm ±0.47 in	±12mm ±0.47 in
	Length Sizing accuracy	±7mm ±0.28 in	±4mm ±0.16 in	±7mm ±0.28 in	±7mm ±0.28 in	±4mm ±0.16 in	±7mm ±0.28 in	±7mm ±0.28 in

Table 1: MagneScan SHRP detection & sizing accuracy

This comparison shows how much the accuracy of an MFL inspection has changed over the past 20+ years.

Factor 1: The Inspection Vehicle

An ILI service starts with a successful run of the inspection vehicle. The design and performance of the vehicle is critical to successful navigation through the pipeline, but perhaps more importantly delivers the ability to detect (POD) defects along the pipeline with enough information to allow the data analysis process to confidently identify (POI) and size (POS) these detected defects.

In 2008 Baker Hughes introduced a new MFL technology system to the industry: the latest generation of MagneScan inspection vehicle. The 6” system launched at the time (Figure F1.1) made use of industry leading electronics and sensing technology to enable step change improvements in sensor spacing, scan pitch and operating parameters. These advancements, and recent others have contributed to the successful roll out of the latest generation MagneScan vehicle to its current capabilities covering 6 – 42” diameter range.

Figure F1.1: Baker Hughes latest generation MagneScan

Previous Baker Hughes reports and publications [ref 1, 2, 3] explain in detail how these vehicle attributes contribute to achieving specifications being delivered today (Table 1). Notably, studies identified that the vehicle alone can only take specification improvements so far.

Specifically, it was found that there is a non-linear relationship between sensor density and signal sizing performance. There is an optimal sensor density above which detection and subsequent sizing performance will not improve significantly, even if the vehicle were to have an ‘infinite’ number of ‘infinitely small’ sensors.

In other words, the inherent physics in the amplitude responses and signal-to-noise thresholds of any real system **do not provide a beneficial improvement of the signal detection or signal characterization with radically improved sensor spacing**.

The physics of MFL signal spatial distributions, the local magnetization levels and signal interpretation, ultimately within the cross-analysis/synthesis process steps, were key considerations resulting as an overall system to maximize feature (e.g. pinholes, slots, pits, etc) detection and sizing entitlement.

Factor 2: Software & Feature Recognition

Specialised software and algorithms are essential to the analysis of pipeline inspection data; they support the analysis process by enabling manual analysis to focus decision making on the regions and features which are most critical and where manual expertise adds the most value (Figure F2.1).

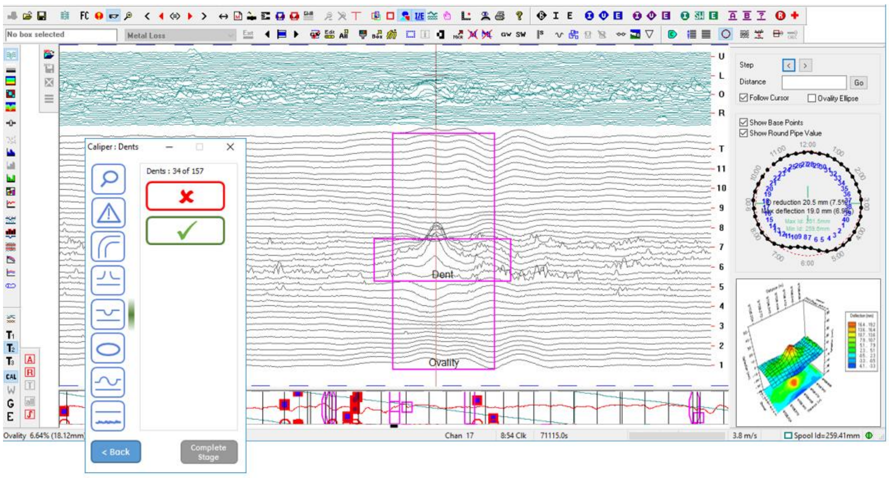

Figure F2.1 The caliper decision support workflow user interface

The signal data collected during an ILI run can be represented as a grid that covers the whole pipeline wall surface and, by analogy, can be thought of as an image of the pipe. For a 100km pipeline section, this image may be 1000 pixels high and 50 million pixels wide (number of sensors x number of data scans), and the task of ILI data analysis is to identify, classify and quantify the size and severity of any injurious features in this massive data set. The number of individual corrosion pits in a pipeline this size may run into millions, and although all of these are visually inspected, algorithms are required to locate and pre-assess this volume of features.

For features meeting the system POD specification to be reported the ILI analysis process must be able to correctly identify and classify them (POI). Achieving a high POI has two components: reliably detecting and labelling areas of data as a region of interest, and then accurately classifying the cause of the signal detected for each area (Figures F2.2 and F2.3).

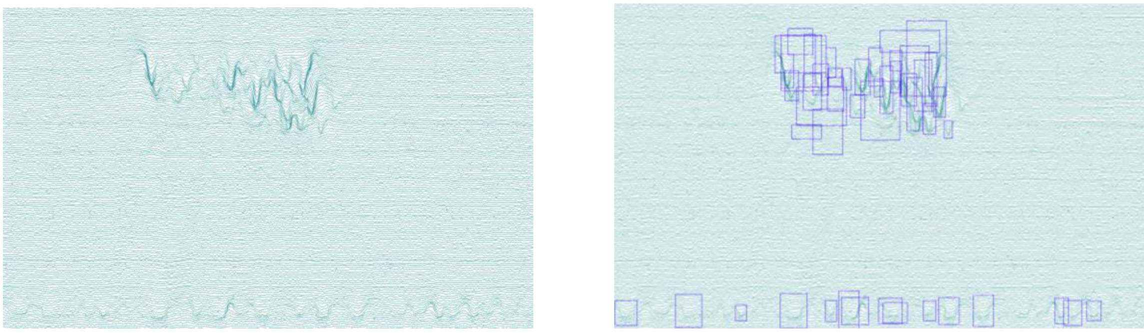

Figure F2.2: ILI signal data showing an area of corrosion above and a seam weld below. On the right we can see potential features detected on both areas by the 'boxing' algorithm.

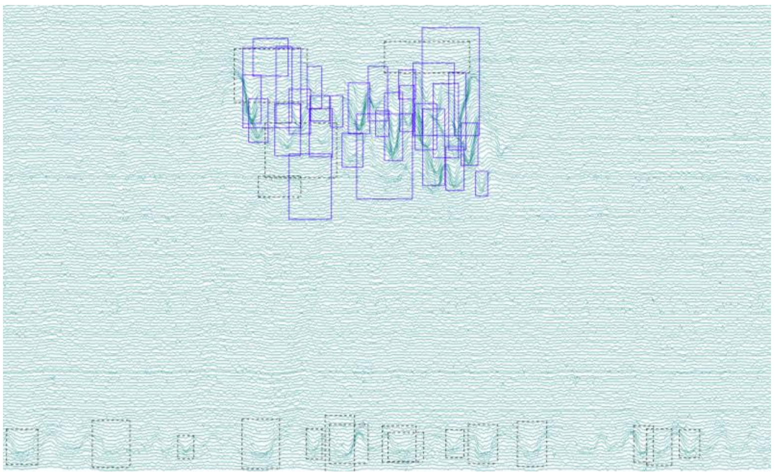

Figure F2.3 After classification the seam weld and non-corrosion areas have been removed

As MFL technologies do not provide a direct measure of defect depth, MagneScan feature detection algorithms ensure that all features meeting the POD specification are detected, from the largest area of general corrosion down to 5mm diameter pinholes. Advanced pre-processing is used on the raw ILI data to normalise, improve signal-to-noise, and ensure consistent detection across all wall thicknesses and pipe types. POD and POI detection specifications are verified for every ILI system by including features into pull tests which are at and below the expected detection thresholds.

The nature of an accurate MFL inspection system is such that it can be very sensitive to variations that are often seen in different pipelines, even if they are considered the ‘same’ (WT, steel grade, corrosion levels etc) on paper. This ‘pipe-to-pipe’ variation is one of the biggest challenges to accurate classification. To overcome this, and meet the accuracy and reliability needed, the latest generation of Baker Hughes classification algorithms are trained and tested on a data set consisting of hundreds of individual pipelines which total around 40,000 Km, contain over 250 million detected metal loss features, and have 100 Terabytes of recorded ILI data.

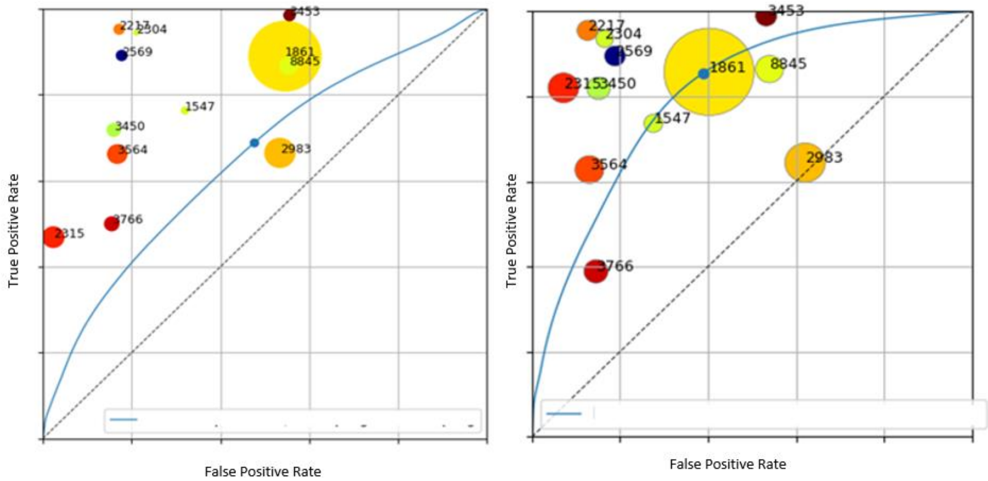

Figure F2.4 Example of iterative performance improvements during algorithm development Each circle represents features from individual pipeline sections with varying attributes.

Development does not stop once the algorithm is being used live in production. Performance metrics built in to the analysis software continue to be gathered with each inspection to measure performance and capture unusual line conditions that are used to update and improve the algorithm over time. The example on figure F2.4 shows the area above the blue line is reducing. A reduction in area above the line represents an increase in accuracy through algorithm refinement.

Factor 3: Data Analysts & Data Analysis Process

Data analysis is where the bulk of the ‘time’ is spent during any pipeline ILI service. Although there is no direct correlation between the time spent analyzing the data and the typical contractual reporting timescales, it is still a good indication of the levels of ‘effort’ required. A typical MFL inspection report timescale is 60 days from receipt of the data to deliver of the report (this time will increase for longer pipelines e.g. 100 days for pipelines >150km or even 140 days for EMAT technologies). Although there are sophisticated feature recognition algorithms and software techniques applied to the ILI data before detailed analysis starts, every inch of the ILI data is reviewed by a data analyst. As this stage is so critical to report quality and resulting end-product accuracy (Figure F3.1) Baker Hughes invests in ensuring the **right people** are selected and governed by **robust processes.**

Figure F3.1. Holistic view of factors influencing ILI report quality.

Data analysis essentially consists of spending many hours a day, often for weeks at a time, looking at a busy computer screen of lines and colours for patterns and 'stand out' features. It's often a case of making sure the software got it right – and it doesn't always! This challenging work takes a certain type of individual, hence, Baker Hughes strive to recruit and retain engineering degree level candidates that go through 'psychometric' screening to ensure they have the right 'minds' for the job. This screening is designed to make sure the candidates have both the necessary attention to detail and the ability to commit to the role for a number of years. The latter is clearly important when you consider the time it takes (Figures F3.2, F3.3) to gain the experience and qualifications necessary to comply with the internationally recognized standards ANSI/ANST ILI-PQ-2017. The full details of how Baker Hughes complies with ILI-PQ-2017 are documented in formal document reference Global-E-M003.

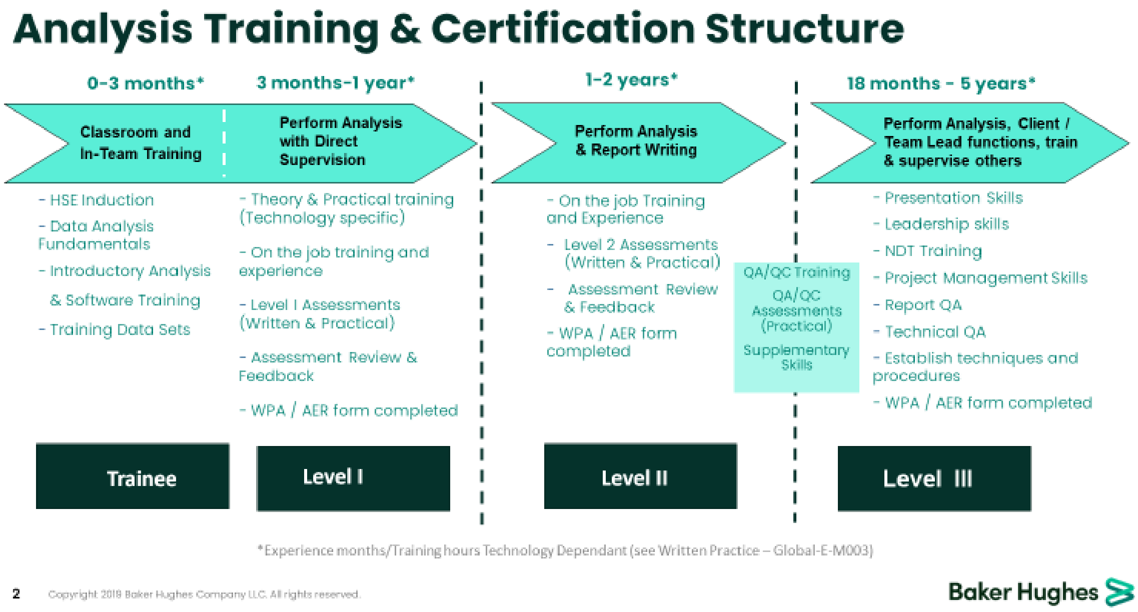

Figure F3.2: Baker Hughes Analysis Training & Certification Structure

Level	Experience (Months)	Training (Hours)	Educational (Formal)
Level I	6	80	*
Level II	18	160	*
Level III	36	500	**

Notes:

* High School graduate or equivalent
** Completion with a passing grade of at least 2 years of engineering or science study at a university, college or technical school.

Figure F3.3: Baker Hughes magnetic technology qualification & certification requirements

The data analysis teams work to, and are governed by, a range of processes and procedures. These are controlled within the ISO 29001 Quality Management Certification system which exists at every one of the Baker Hughes operational sites. Notable elements contributing to robust processes and procedures include:

- On the job training (OJT)
- Report Audits
- Continuous Improvement & Feedback system
- Data analysis quality metrics

These, and other elements, are covered in more detail in recent publications (reference 4), but it is worth exploring **report audits** in more detail.

Internal post-delivery report audits are an important best practice. These provide a means of making sure analysis and reporting standards continue to meet the stringent high-quality requirements expected by customers and that significantly impact report accuracy. All audits should be planned, documented and scored to provide the foundation for generating ongoing analysis quality metrics. This proactive approach should seek out potential errors and highlight any process issues that have the potential to introduce future error. Should an error be found, it is documented, which initiates a formal Root Cause Analysis (RCA) and corrective action is taken.

Factor 4: Algorithms & Sizing Models

It is not enough to detect an area of corrosion and report it as such, an ILI inspection also needs to report the depth and extent of that corrosion accurately. The level of sizing accuracy that can be achieved (POS) is usually stated as a tolerance +/- a given percentage of the pipe wall thickness, and calculated to an 80% or 90% confidence level, meaning 80% or 90% of all corrosion features will be expected to meet the given tolerance.

The task of predicting the depth profile of an area of corrosion is not straightforward. The relationship between the recorded magnetic flux leakage and defect depth is complex and highly nonlinear; even for simple isolated pits sources of variation include the ILI vehicle build, magnet strength, wall thickness, pipe material, vehicle speed, and of course, the shape of the pit itself.

The Baker Hughes process of sizing consists of two aspects; first characterising an area of corrosion using several descriptors, and then using those descriptors to predict the corrosion dimensions using a statistical method called a 'sizing model'.

Sizing models start with a carefully chosen population of artificial defects machined to replicate real corrosion. 'Pull Through' tests are carried with every ILI vehicle on these defects to give comprehensive coverage over all defect shapes, wall thicknesses and speeds; the sizing model is built using this data.

Figure F4.1 Pull through pipe spools with machined defects

The introduction of the latest generation MagneScan fleet in 2008 saw a step change in the defect population size and variation, resulting in an improved POS across all defect morphologies. Sizing models are now typically derived on an extensive range and number of individual defect signals, and crucially incorporate the expertise and knowledge accumulated across decades of experience in ILI inspection to create a model that is robust and accurate across the whole population.

Although they share the same form, each sizing model is uniquely tailored to an ILI system configuration to ensure the best performance. This means that Baker Hughes has created over 500 models to date.

The POS performance is measured across all defect shape categories in the pull through data set, and due to the variation and extreme defects in this population it is often found that the model performance in operational data, where the natural corrosion is more typical, will exceed the stated POS.

Factor 5: Performance Validation, Verification & Improvement

The first part of the 'proof' of performance of an MFL inspection system is validation using 'pull-through' data. This compares the recorded, analyzed and sized signals vs the known actual defect dimensions in the pull through spools. Each new MFL vehicle design in the Baker Hughes fleet goes through this validation prior to its release into operations. As mentioned earlier in this paper the first of the latest generation MagneScan fleet was the 6" vehicle - its performance validation can be seen in figure F5.1. In this case the results proved that the vehicle exceeded the depth sizing accuracy target of +/-10% WT with 90% certainty.

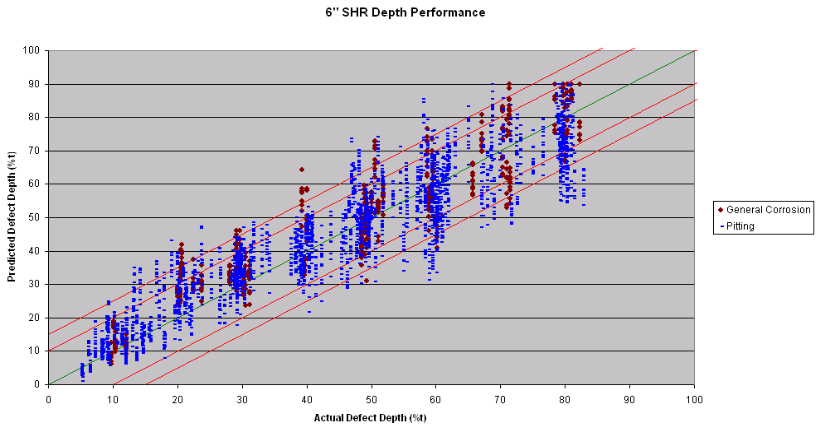

Figure F5.1 Validation of the 6" MagneScan system

Following the system validation and operational release, it is then critical to operator confidence that this can be followed up in the field. Below, figure F5.2 shows how sizing performance of the MagneScan system was verified from multiple sets of dig data provided by operators in Asia, Europe and North America. The system is consistently performing at greater than 90% certainty.

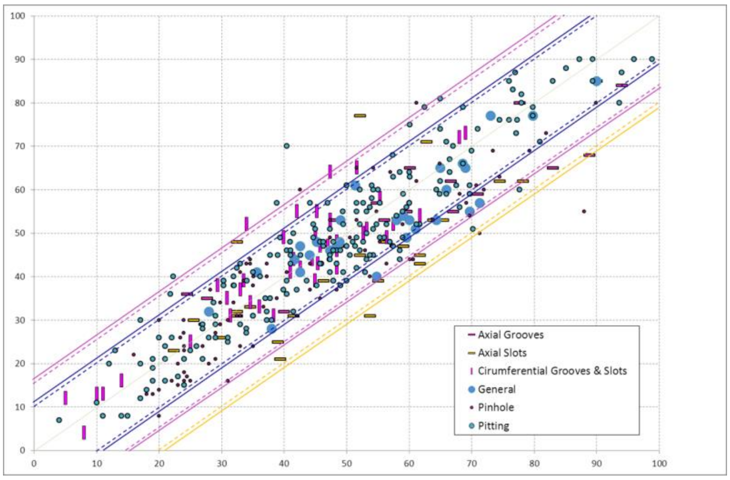

Figure F5.2 MagneScan dig verification data unity plot

As the volume of 'truth data' grows, confidence in the accuracy of the system does also. In parallel, opportunities for improvement are also presented. In the case of the MagneScan system, a significant improvement opportunity arose to expand the range of features that could be detected, identified and sized accurately. This improvement is covered in detail in an earlier publication (reference 3) but it led to the release of the MagneScan Super High Resolution Plus (SHRP) specification which added detection and sizing accuracy for pinholes and slots. Figure F5.3 shows the performance of the MagneScan SHRP with respect to pinholes within areas of general corrosion.

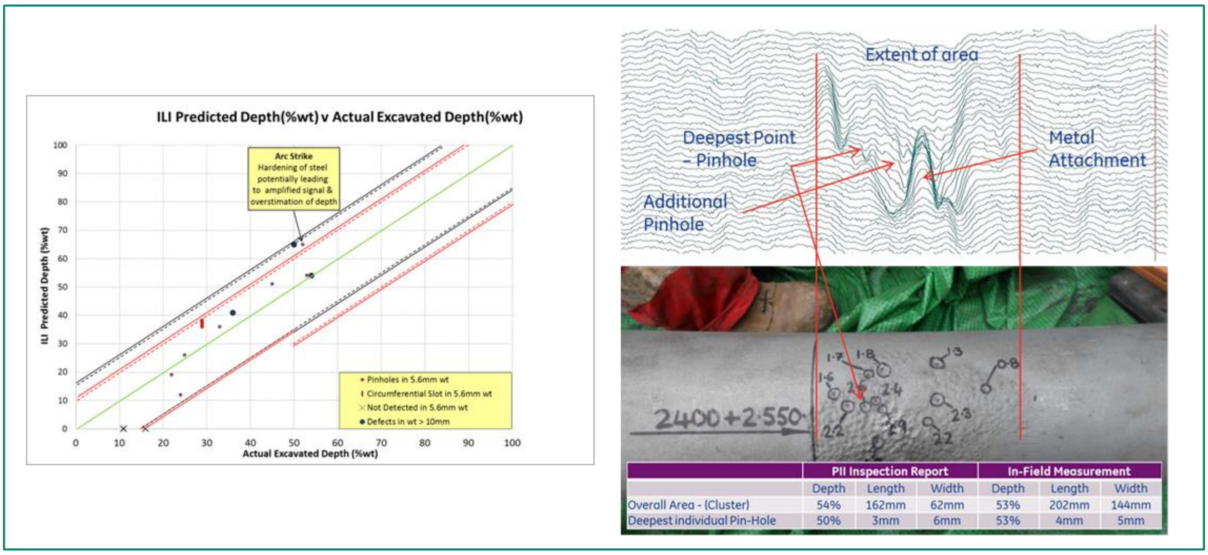

Figure F5.3 MagneScan pinhole verification

Since laser scanners have become the norm when verifying ILI performance, the Baker Hughes dig verification data base has grown exponentially from a few thousand defects prior to 2015 to hundreds of thousands today, across all 7 POF categories. Matching of laser scan excavation data is carried out using the *DigCom* software introduced in 2013, this software allows matching of each individual pit even in complex corrosion (Figure F5.4).

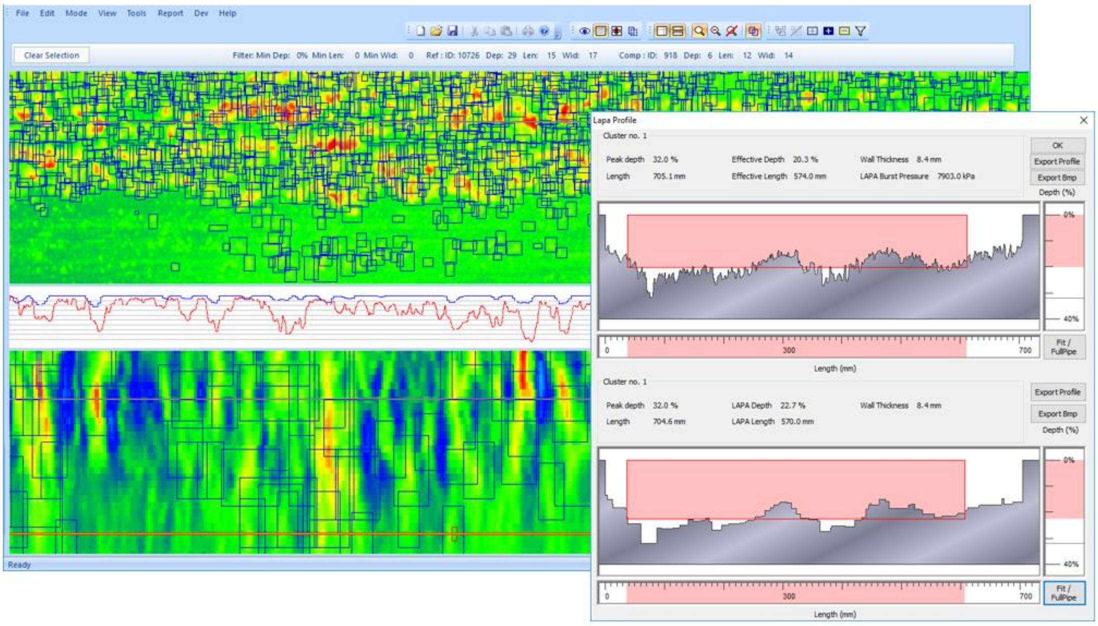

Figure F5.4 Matching laserscan (upper) and ILI (lower) data in the DigCom software and the resulting riverbottom profile from laser scan compared to the ILI 'boxed' data

The significant growth in truth data has led to Baker Hughes introducing regular accuracy performance reviews. Held quarterly within our organization and annually with many key customers, these reviews allow us to consider results in detail with the aim of continually improving our accuracy and overall offering to operators. The current truth database for the latest generation of the MagneScan fleet contains in excess of 60,000 features reported at the most accurate (SHR/SHRP) specifications. Actual performance is proven to significantly exceed stated specifications of POD, POI & POS @ 90%.

Since the introduction of these regular reviews, trends and early indicators are being used to drive multiple improvement and enhancement initiatives such as (but not limited to):

- Defect outlier elimination
- Girth weld crack detection & sizing (reference 5)
- Automatic prediction enhancement
- Training and processes

Conclusions

As noted in the introduction (and discussed in greater detail by Bluck, Sutherland, Dawson [ref 6]) ILI accuracy plays a significant role in achieving critical assessments of pipelines. This accuracy has a direct influence on both:

- **material cost saving of reduced digs;** and
- **improved pipeline safety**.

This paper has identified the main protagonists that contribute to the delivery of reliable & accurate data supplied by an MFL, or indeed any ILI, inspection service.

Whilst the ILI vehicle often takes center stage it is supported in equal measure by several other factors. It has been shown that as far as the vehicle is concerned 'more' doesn't necessarily mean 'better' or specifically 'better accuracy'.

At Baker Hughes there is a belief that, based on current industry hardware, accuracy improvements that can directly influence critical assessments of pipelines are just as likely to come from what we do with the data we have today as they are from improvements on the vehicle itself.

Acknowledgements

The authors wish to thank the various members of Baker Hughes who have provided input and previously written papers that have been used to support this paper.

References

1. *Pinhole Defects Detectability: Investigation into the smallest defect detectable with MFL technology*, Mullin, I., Canni, G.

2. *Validation of latest generation MFL in-line inspection technology leads to improved detection and sizing specification for pinholes pitting, axial grooving and axial slotting*, Submission IPC2010-31124, Sutherland, J., Bluck, M., Pearce, J., Quick, E.

3. *Advancements In The Detection & Sizing Of "Pinhole" Metal Loss,* Submission IPC2014-33128, Santamaria, W., Bluck, M.

4. *In-line inspection: Both effective data collection and interpretation needed to achieve high quality reporting results*, Submission IPC2010-31112, Walker, J., Mallaburn, P., Balmer, D.

5. *Advancements in the identification and inspection of Girth Weld Anomalies,* Submission IPC 2012-90114, Bluck, M., Sutherland, J., Pearce, J.

6. *Achieving critical assessments of pipelines through accurate & reliable inspection information*, Submission NACE 2nd Conference & Expo Genoa 2018, Bluck, M., Dawson, J, Sutherland, J.