Delivering customer impact with data-driven product development

Philip English,
Baker Hughes Process & Pipeline Services

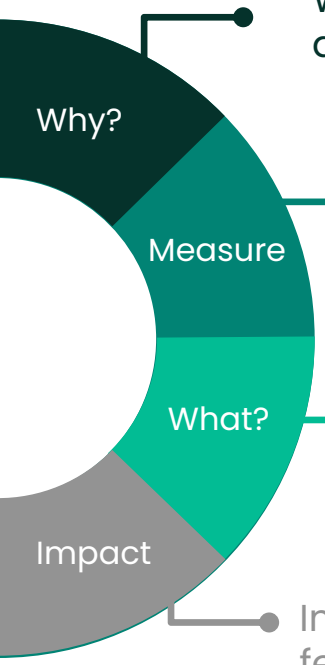
Co-author
Anthony Tindall
Baker Hughes Process & Pipeline Services

PPSA 2025

Agenda

1. Understanding the Limitations of a Simple Specification and Integrity Management
2. Artificial Intelligence - Collecting the Right Data & Benefits of Triaxial MFL
3. Validation - Building a Comprehensive Dig Database
4. Generating “True Defect Profiles” from Triaxial MFL Data
5. Defect Specific Individual Tolerances
6. RunCom 3D- Statistical Dynamic Growth
7. Artificial Intelligence- Measurable Impact

Intentional development



What is the problem or improvement to be addressed? Why is this good for the customer?

Evaluate current data sources.
What do we have? What do we need?

Identify and execute specific projects.
Retain focus on problem statement.

Impact needs to be measurable, meaningful and felt by customer.

The in-line inspection industry has measured success via:

1. Pipeline Operators Forum defect classifications, in conjunction with
2. API 1163 definitions on how to measure statistical population performance

These methods pose challenges to pipeline operators when real life situations occur.

Let us consider the following two familiar questions:

"The corrosion on my pipeline does not fall into the simplified 7 POF categories; how do I measure performance on complex defects like pit in pit and other interacting defects?"

"It is great that the ILI tool was within specification on the 98 features from infield verification on my pipeline, but I have 2 outliers, not in specification which nearly caused a leak. How do I manage this effectively?"

Baker Hughes

Why?

- Simplified specifications are often based on idealised parabolic shapes

- While this simplified approach is widely used, it becomes inadequate when dealing with complex corrosion features:
  - defects within defects
  - intersecting axial slots
  - corrosion surrounded by background material

  Failing to account for the full range of real-world corrosion scenarios.

Pipeline operators dealing with real life complex corrosion features can often face two main challenges:

1. Safety Outliers: These occur when the true severity of a defect is underestimated, potentially leading to catastrophic failures like leaks or ruptures.

2. Resource Outliers: These arise when the severity of a defect is overestimated, leading to unnecessary excavations and repairs.

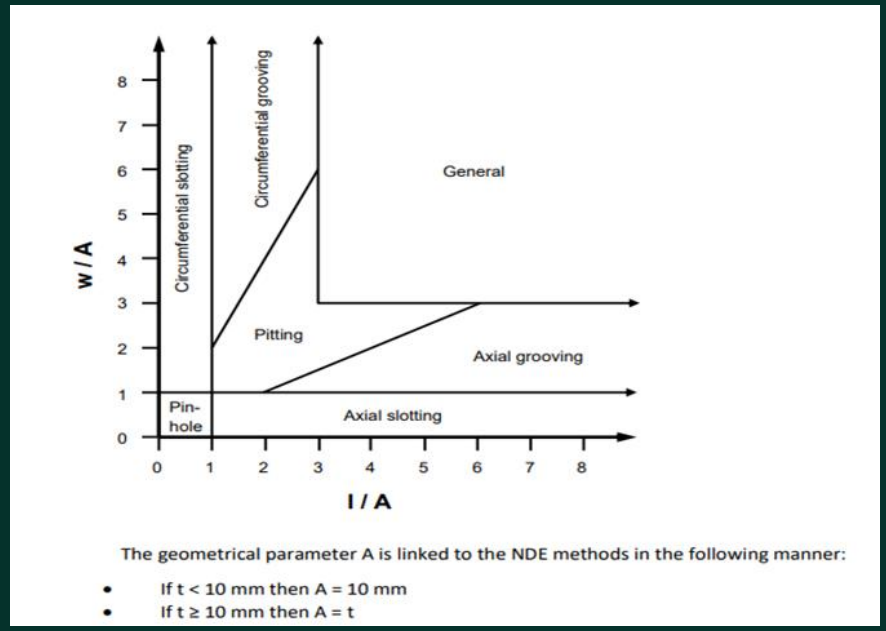


Figure 1: POF Metal Loss Categories

Simply applying a static specification to a defect POF category limits the usefulness.

Consider a simple specification for +/-10% for all pitting; even those untrained in Magnetic data analysis can see that applying this to a 10% defect and a 70% does not logically fit.

This is a simple example and if you consider interacting defects compared to isolated features, this concept of a static specification applied for all pits rapidly becomes illogical.

Baker Hughes

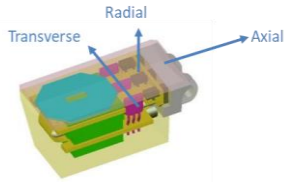
Collecting the Right Data

Data-driven developments are useless without the right data

Pedigree past, present, and future

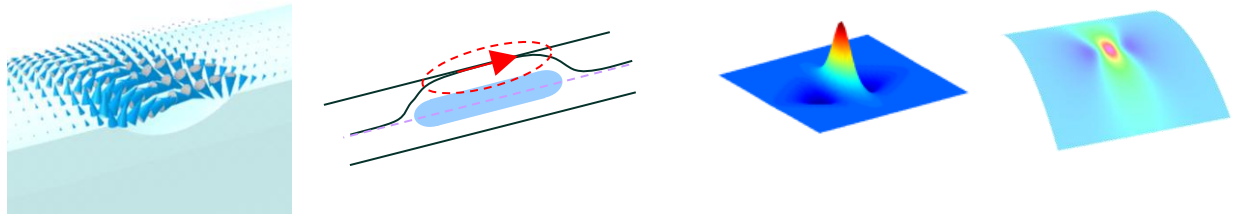
Baker Hughes

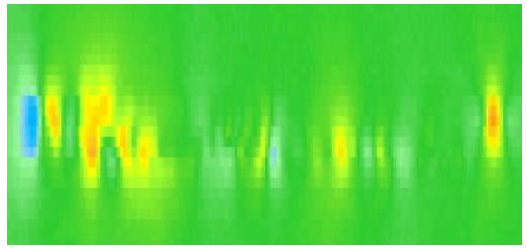
Collecting the Right Data - Triaxial



Axial

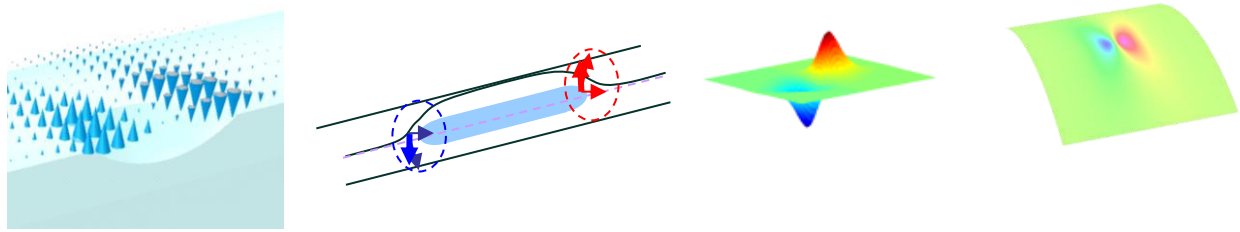
- Measured down pipe
- Sensitive to volumetric metal loss
- Measure nominal field strength

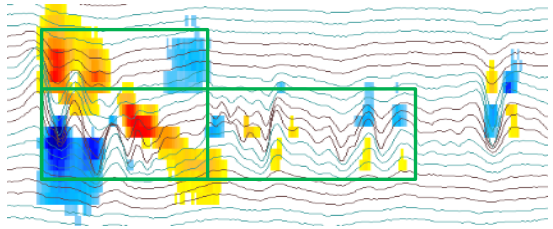




Radial

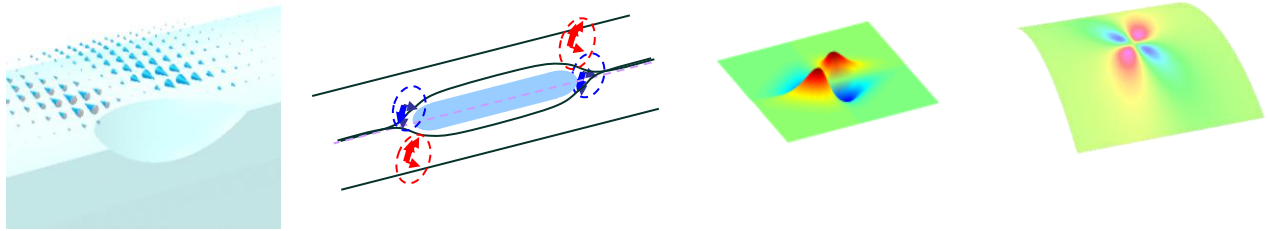
- Measured out/in from centre
- Highest sensitivity to changes in depth
- Identifies start and end of features

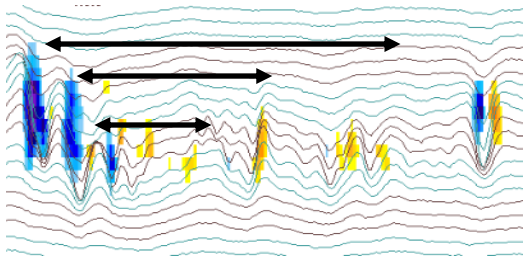




Transverse

- Measured circumferentially around pipe
- Added sensitivity to in-plane shape
- Improved width measures & ID of background corrosion

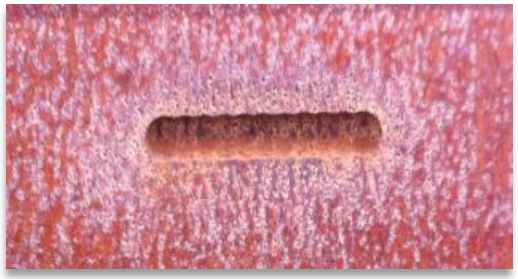
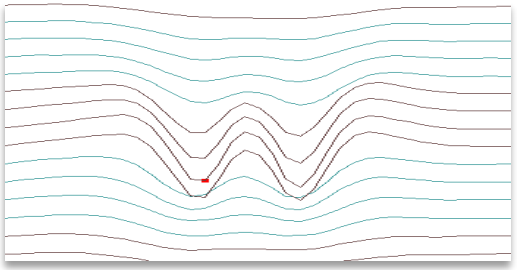
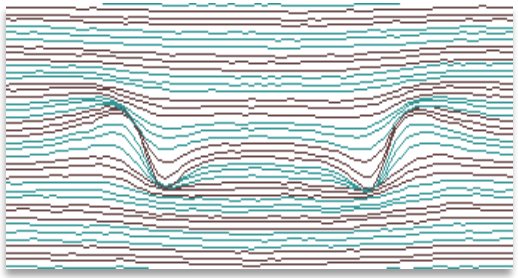
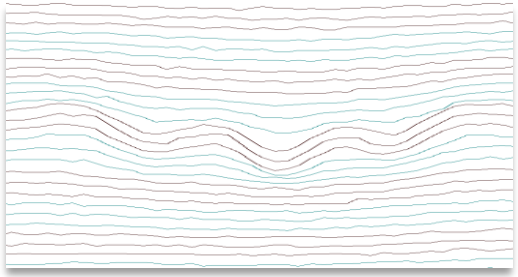
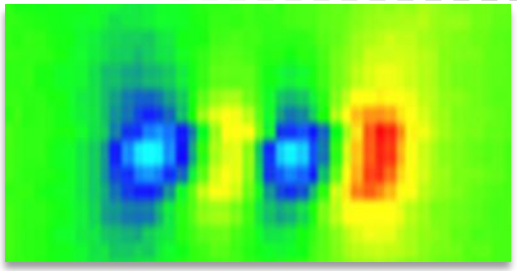
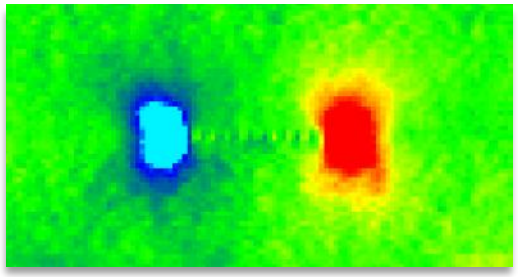
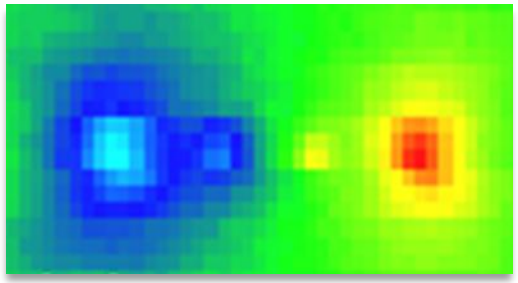




Three views of each defect.

Baker Hughes

Triaxial Interaction – Basic Principles

	Interacting Pits	Axial Groove	Defect-within-defect
			
Axial			
Radial			
	Covered by POF Classification specification		NOT Covered by POF Classification specification

Triaxial data allows for the distinction between close interacting pits and axial features.

Defects within defects can easily be interpreted as three pits on axial only inspections.

However, with the Radial signal it's clear in the signal:

1) Large start (Blue)
2) Small Start (Blue)
3) Small End (Orange)
4) Large End (Red)

= slot in slot

The radial signal clearly differentiates between corrosion morphologies

Baker Hughes

Validation - Dig Verification Database

Quality in Quantity is key

Transforming the future of corrosion management

Baker Hughes

Next Generation MFL Sizing

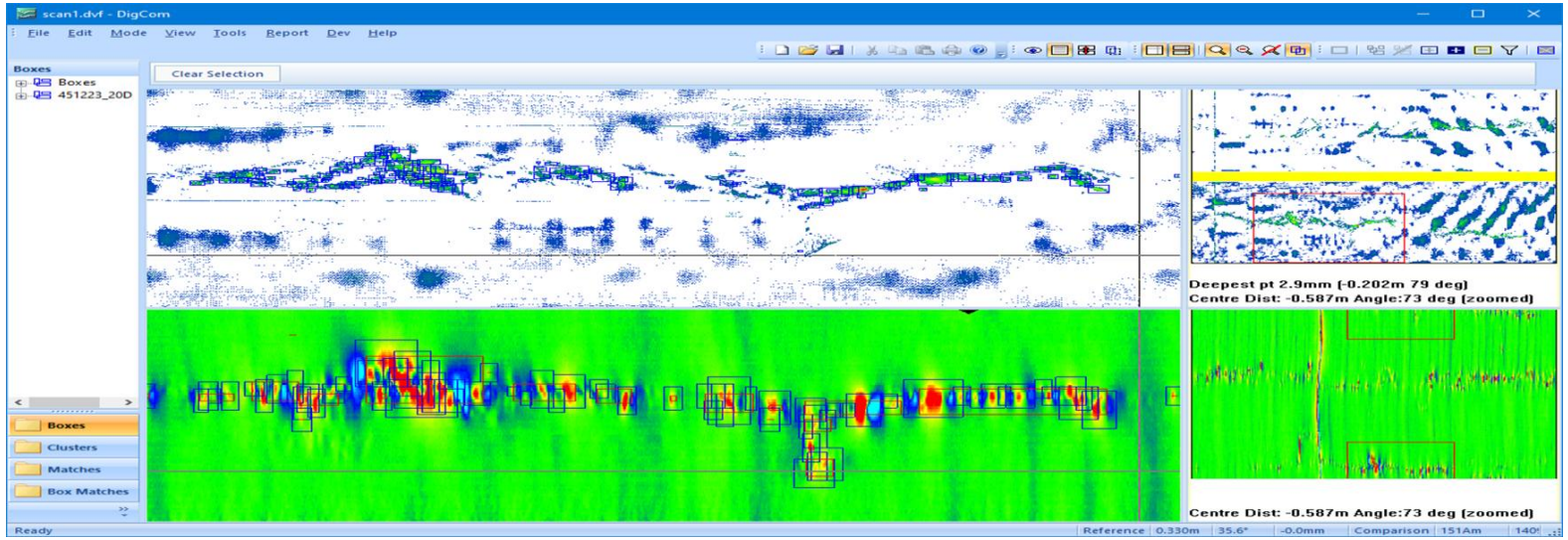
Data Collection:

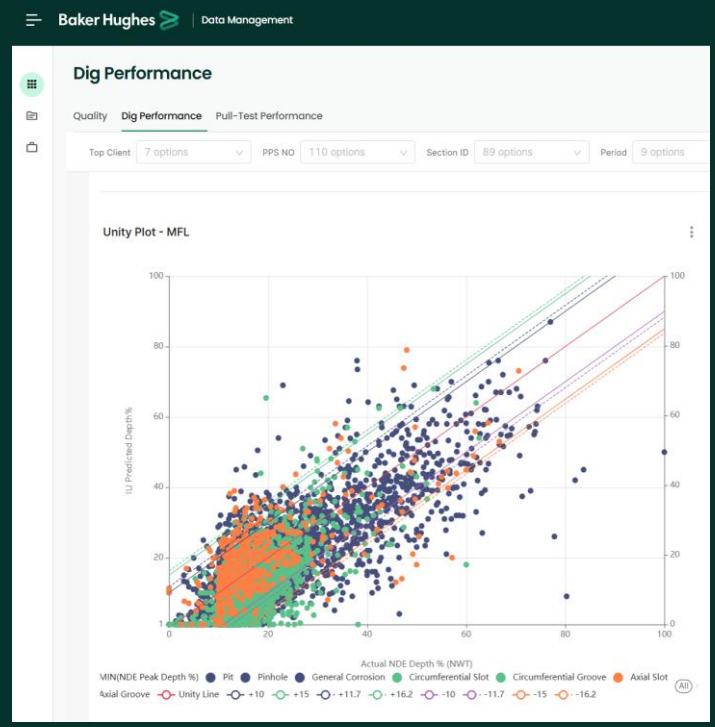
- Select for high reliability data, well conditioned for matching to ILI data.

Data Ingestion

- Import matched ILI and dig data with associated metadata.

- Over 150,000 excavations collected to date
- Each 2-6m in length
- Over 400,000 defects used for model generation
- Exponential increase due to introduction of Laserscan.





The goal of the Dig Verification database is to:

1. Monitor performance and trends
2. Share information with clients
3. Training data for statistical models

Due to its completeness, ideal as the truth data needed to train machine learning models

Baker Hughes

Next Generation Sizing

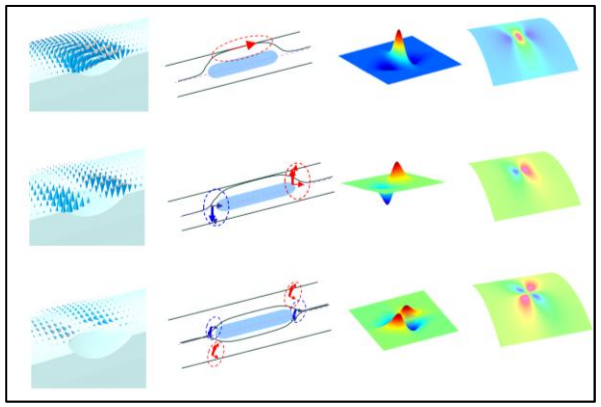
Accuracy enhancements to MFL analysis

Baker Hughes

Next Generation MFL Sizing

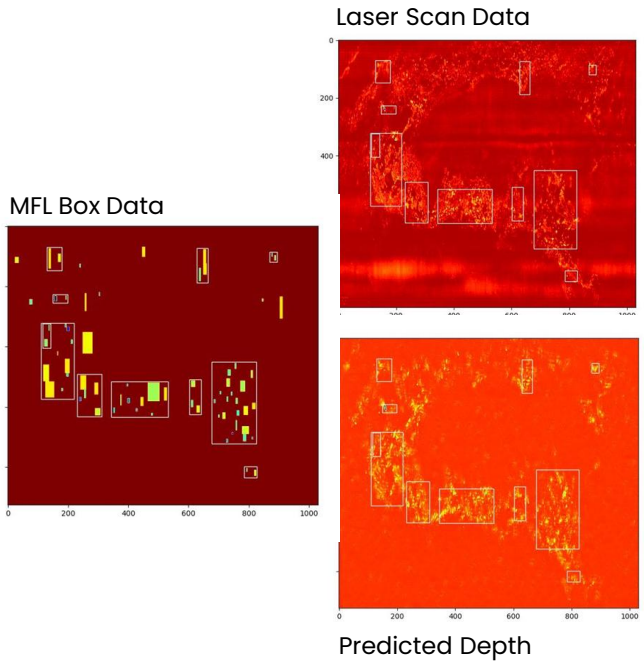
Customer Value

- Material improvements in dig effectiveness and efficiency
- Targeting step change in improvements to outlier performance
- Improved confidence in accuracy predictability – reduced safety risk
- Significant report cycle time reduction

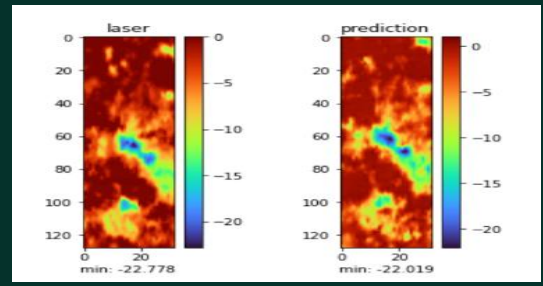


How is it different?

- Triaxial data as a foundation
  - 3 looks at the data for model training with 1 technology/run
  - No data alignment challenges – perfectly aligned
  - Multiple runs in segments
- Deep learning techniques based on unique training sets
  - No simulated data
  - Database of over 400,000 aligned laser scan feature correlations
  - Multiple diameters, conditions, products
- Data processing performance
  - Parallel cloud processing techniques developed with Baker Hughes CET group proven to allow step change in prediction times



FAST FACTS



Deep learning expected to be initially used to augment current analysis by end of 2025

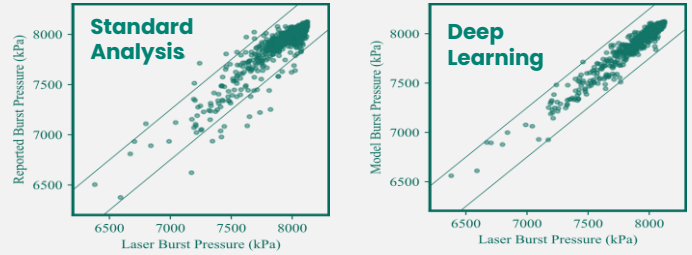
IPC Paper to launch capability

Technical development completed with excellent results for depth and burst pressure prediction

Major investment in Data Comparison development tool-set (catalogued & aligned prediction and laser data for all database defects)

Corrosion depth and pressures demonstrated on highly corroded pipe (whole line) in a few hours

Improved burst pressure prediction
Reduced Outliers



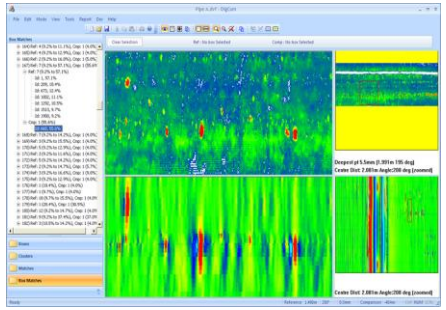
Defect Specific Tolerances

Removing conservatisms

MFL Predicted Tolerances

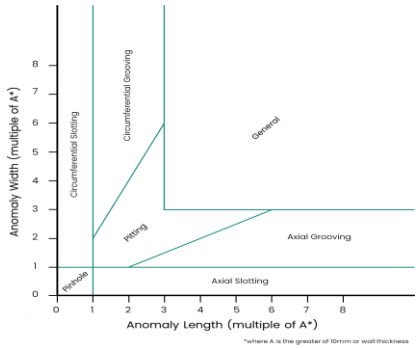
Customer Value

- Tolerances express uncertainty in predictions
- Including tolerances in defect assessment reduces risk of safety outliers
- Established tolerance structure does not reflect reality of inspection accuracy factors
- Improved tolerance accuracy can help improve effectiveness of dig programs



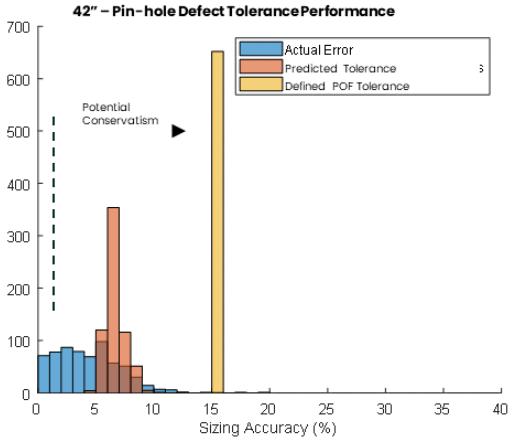
POF Based Tolerances

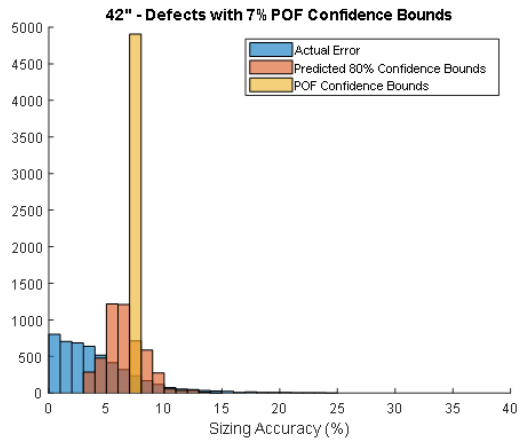
- Corrosion tolerances built on arbitrary POF based length and width categories
- Defined with limited defect populations, often using manufactured defects



Anomaly Specific Tolerances

- A machine learned model
- Over 500 MFL data attributes considered, utilizing >100,000 laser scan dig results





- Consider the simple question:
- *"Will a pinhole be more accurately sized if it is isolated within a spool or within a large area of corrosion?"*
- Therefore, other factors need to be considered:
  - Depth
  - Is the defect within another defect (e.g., pit in axial slot)
  - Neighboring defects
  - Number of neighboring defects
  - Shape of neighboring defects
  - Depth of neighboring defects
  - Proximity to girth weld
  - Proximity to wall thickness changes
  - Proximity to field changes (casings, sleeves, etc.)

FAST FACTS

Demonstrated potential to reduce digs by up to 62% and provide potential to extend re-inspection interval while maintaining safety

Collaboration and feedback with customers

Baker Hughes

RunCom 3D

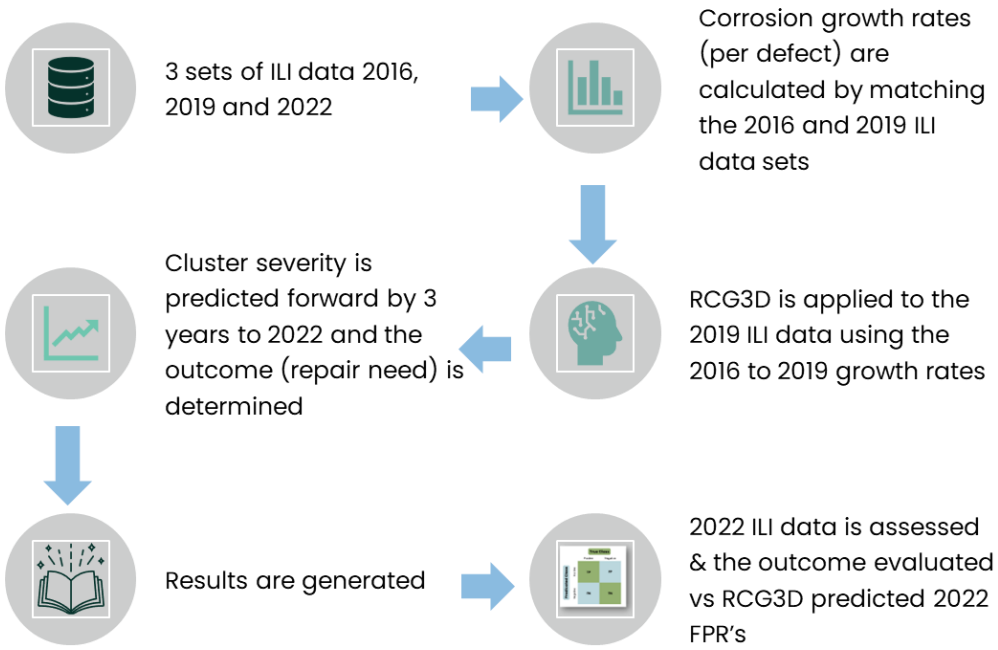
Improved prediction of future threats

Baker Hughes

Runcom Cluster Growth 3D (RCG3D) Assessment

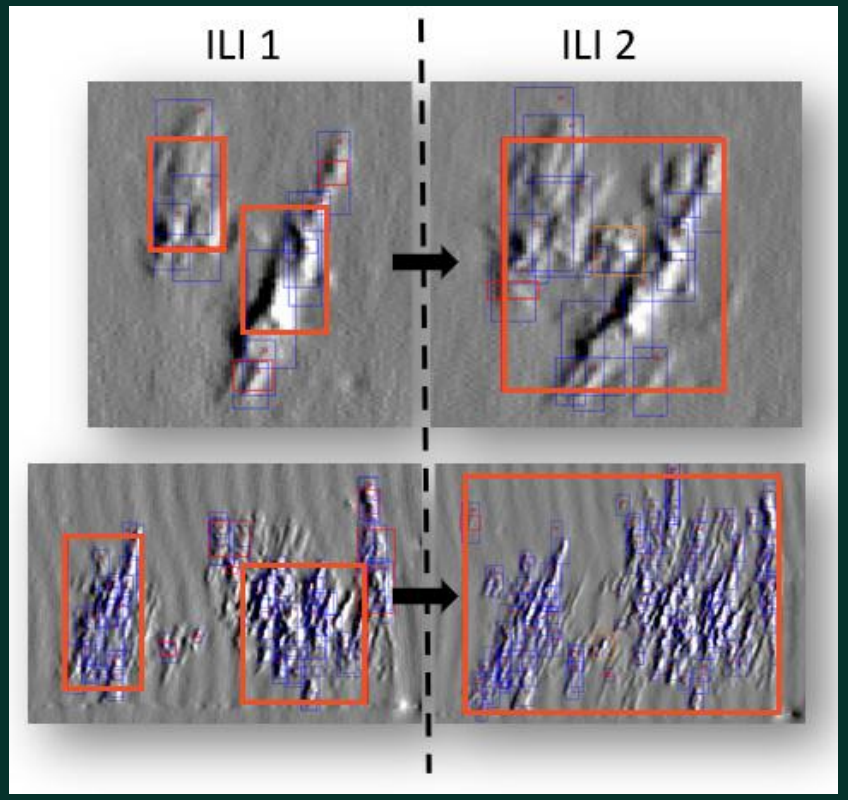
- A new approach to predict corrosion development over time

- Uses local corrosion growth rate distributions, applied using a probabilistic methodology
- Accounts for growth in depth, length and width dimensions
- Includes prediction of the location and pattern of potential new corrosion sites using a machine learning model
- Considers the interaction between existing & predicted new sites over time
- Improves response time predictions and minimises unnecessary digs



3D Growth Assessment

Areas of corrosion which are reported by ILI data analysis as "clusters" will change over time in all 3 dimensions (depth, length & width dimensions) as the existing sites grow and new corrosion sites occur in and around the edges of existing areas



RCG3D models the potential for future interaction between nearby areas of corrosion via a machine learning neural net model that predicts plausible positions for new defect initiation

Baker Hughes

Measurable Impact

Quantifying customer experience

More than just Smart

Baker Hughes

Data Accuracy Metric – Criteria and Validation

Objective

- Better model customer's subjective experience.
- Continue objective drive to eliminate pipeline failures.

New Metric – Data Accuracy Score

- Enables PPS to proactively and confidently engage with a customer on ILI performance
- Lagging indicator acting as "Red Flag" for early warning of:
  - When a customer **may be dissatisfied** with performance
  - When we **may have a quality issue**
- Positive and negative performance engagement

Performance to Expectation

Calculates 2 measures for a pipeline segment

Performance to published specification

POD / POI / POS to API1163

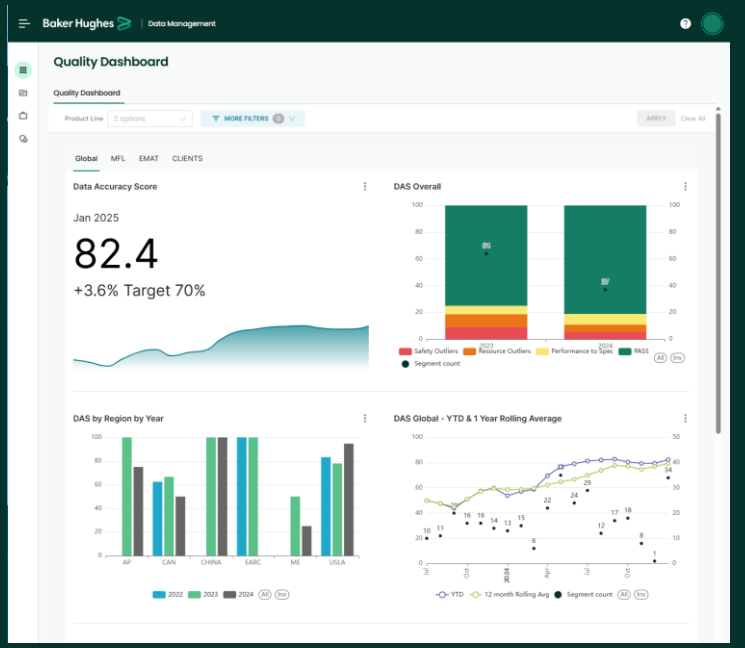
+

Outlier performance definition

Resource or Safety

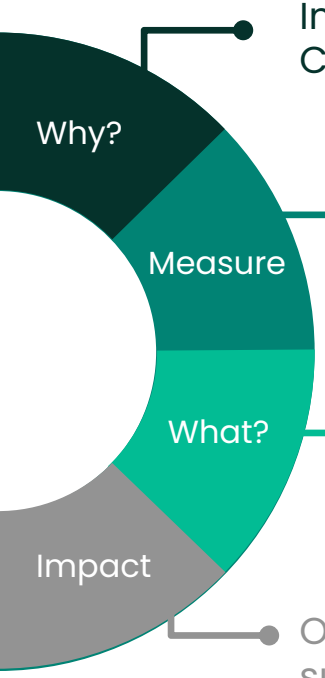
Acceptance criteria

Combination of safety outliers, resource outliers and performance to specification
PASS only if all 3 conditions are met



Data Accuracy Score Dashboard
(In progress of ingesting past data)

Intentional development



Improve customer value and experience.
Continue to drive down pipeline failures.

Maximise value from triaxial data.
Extend and exploit historical dig verification data.

Improved sizing of complex defects, reducing outliers.
Defect specific tolerances, targeted conservatism.
Extended corrosion growth, improved integrity predictions.

Objective safety measures combined with subjective customer experience metrics.

Improve safety

Predictive models, trained on high volume and accuracy field measurements and operational data, are more able to consider the many influencing factors at play that can impact metal loss sizing.

Reduce unnecessary digs

Predictive models that reduce conservative POF-based tolerances and improve depth and burst pressure accuracy mean you only dig what's really necessary

A second opinion

AI-driven assistance provides analysts with an independent view of reconstructed corrosion profiles to support their decision making. Complex corrosion morphologies are the primary cause of under conservatism experienced

Take action sooner

Rigorous full line analysis, powered through multiple parallel cloud based virtual computers gets you what you need earlier to facilitate repair programing. Reduce finalized priority digs from up to 120 days to 10-15 days

Baker Hughes

Baker Hughes