

# Novel Signatures from Deployed Transmission Infrastructure Sensors

Project Number (72954)

Kayte Denslow, William S. Rosenthal  
Pacific Northwest National Laboratory

---

U.S. Department of Energy  
National Energy Technology Laboratory  
2021 Carbon Management and Oil and Gas Research Project Review Meeting  
August 2021

# Presentation Outline

---

## For this session:

1. Introduction
2. Technical Status
3. Accomplishments to-date
4. Lessons Learned
  - Research gaps/challenges
  - Unanticipated research difficulties
  - Technical disappointments
  - Changes that should be made next time
5. Synergy Opportunities
6. Project Summary
  - Key Findings
  - Next Steps

## For reference, as required:

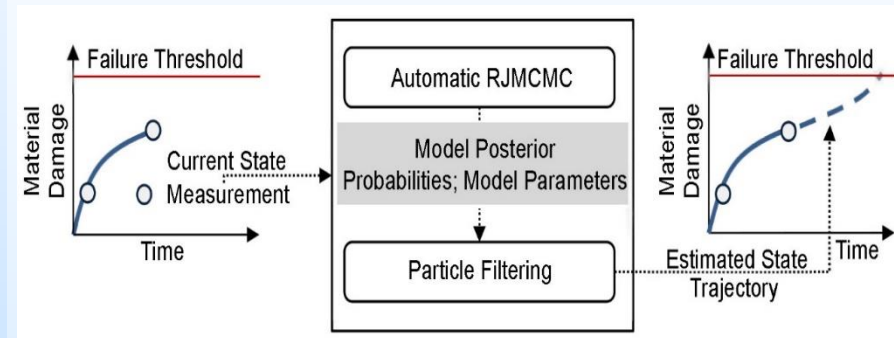
7. Appendix
  - 1. Benefit to the Program
  - 2. Project Overview
  - 3. Organization Chart
  - 4. Gantt Chart
  - 5. Bibliography

# Introduction

## Problem:

- Despite advances in in-line inspection (ILI) tools, unforeseen incidents continue to occur in natural gas transmission (NGT) pipelines that undermine the safety & reliability of the "gas grid"
- **Ultimate Goal:** Use past (existing) ILI data on NGT pipelines to gain new insights on corrosion characterization, corrosion rates, & corrosion driving factors that help operators improve NGT pipeline condition assessment and NGT incident prediction & prevention

1. Improve NGT pipeline corrosion flaw characterization (**diagnostics**) through model-assisted interpretation of ILI signal data to ultimately help operators improve certainty of NGT pipeline failure pressure calcs.
2. Uncover non-obvious corrosion factors to manage for mitigating pipeline corrosion rate
3. Improve certainty of NGT pipeline corrosion initiation time, corrosion rate, and time to critical flaw size to ultimately help operators improve certainty of remaining useful life calculations (**prognostics**) and risk assessments



# Technical Status

---

- Pipeline Reliability Assessment Process
  - Research needs
- Diagnostic and Prognostic Models
  - Description of each
  - Development progress for each
- Envisioned Use

# Pipeline Reliability Assessment Process

Technical Status  
Accomplishments  
Lessons Learned  
Synergy Opportunities  
Project Summary  
Appendix

## Operationally:

### Early detection provides time for preventative maintenance

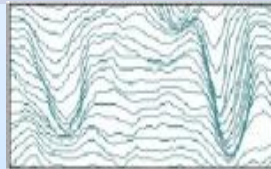
#### Inspection

Collect MFL/UT signal data during in-line inspection (ILI) of in-service pipelines



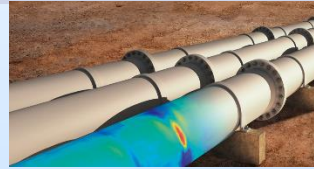
#### Signal Analysis

Process and analyze MFL/UT signal data to derive flaw characteristics (type, size, location, etc.)



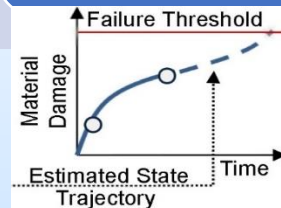
#### Diagnostics

Use flaw characteristics to calculate a pipeline's current failure pressure, determine its fitness-for-service/reliability, MAOP and risk level



#### Prognostics

Use pipeline flaw characteristics with corrosion rate models and failure models/statistical models to predict remaining service life



#### Decision-making

Integrate pipeline condition/risk level and service life forecasts with geospatial models to support planning for NGT pipeline repair/replace/upgrade

**Prioritize preventative maintenance and mitigate failures**

### Research needs:

- *Earlier & reliable* detection of flaws (e.g., corrosion)
- Higher resolution flaw sizing – **novel signatures & data-driven/physics-informed** flaw reconstruction
- Automated analysis of ILI “big” data; *on-line* real-time monitoring (inspected pipelines)
- **Data-driven statistical models** to infer likely condition of *un-inspected* pipelines

### Research needs:

- More certain condition-based predictions of remaining life – **data-driven/physics-informed** models of corrosion initiation times and rates
- **Data-driven statistical models** of remaining life for *inspected and un-inspected* pipelines
- Geospatial map of health/reliability index

# Our Project Focus

Technical Status  
Accomplishments  
Lessons Learned  
Synergy Opportunities  
Project Summary  
Appendix

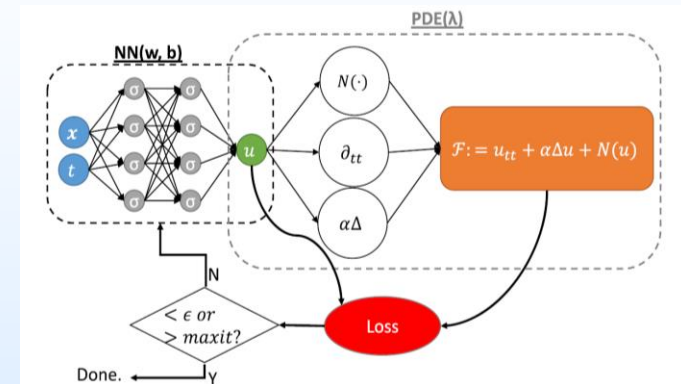
## 2 Models: Diagnostic & Prognostic

### 1. Hybrid data-driven/physics-based model to improve **diagnostics**

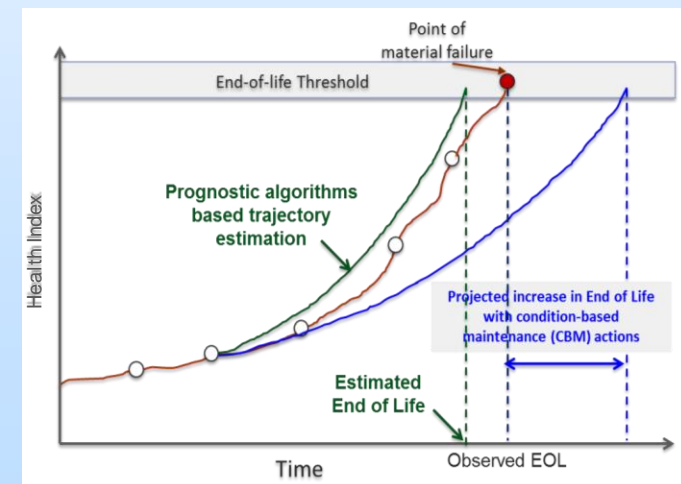
- Being trained to recognize signal features associated with corrosion flaws in magnetic flux leakage (MFL) ILI data
- Being trained to determine physical flaw characteristics (size, profile) associated with the signal features
- Results will support pipeline failure pressure calcs, reliability assessments

### 2. Hybrid data-driven/physics-based model to improve **prognostics**

- Being trained to associate corrosion initiation and growth rate with the pipeline material/environment/construction factors that influence them
- Results will support remaining life calcs



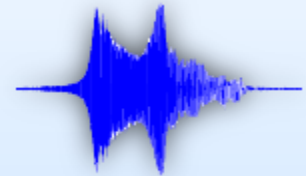
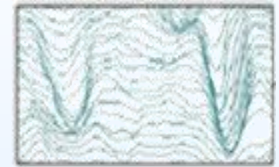
Jagtap AD, K Kawaguchi, and GE Karniadakis. 2020. "Adaptive activation functions accelerate convergence in deep and physics-informed neural networks." *Journal of Computational Physics* 404



# Diagnostic Model

## Corrosion recognition, reconstruction, characterization

- Apply machine learning to large set of MFL sensor signal data collected during in-line inspection (ILI) of NGT pipelines
  - During research projects or real inspections
  - Flaw type focus so far: external corrosion
- Develop model of relationships between physical features of flaws and MFL signal features
  - Raw and/or processed MFL signals on flawed & flaw-free pipelines
  - *And* the corresponding ground-truth data = drawings or confirmatory NDE – e.g., laser profiling or UT
  - Supplement with validated simulations, if practical
- Once tested and validated, the diagnostic model can be applied to other ILI signal data to confidently predict the corresponding flaw characteristics – where ground-truth is not needed
- Applies to *inspected* pipelines only



### Status:




- Received MFL signal data and corresponding ground-truth flaw dimension data (Sept. 2020-May 2021)
- Diagnostic model training is in-progress

# Diagnostic Model Progress





## Phase I (current project scope): Development Phase Between PNNL and Pipeline ILI Service Provider(s)

FY 2019 - FY 2020

Expected Key Outcomes:

- 1)  **Framework for machine learning**
- 2)  **Strategic partnership(s) with NGT pipeline operator(s)**
- 3)  **Sharing of NGT pipeline data sets for a representative sampling of NGT pipelines located across the U.S. needed to train robust machine learning algorithms**

Primary Tasks:


-  Build framework for data-driven model, starting with corrosion
-  Develop partnerships with ILI service provider(s) who have interest in applying machine learning (ML) to historical pipeline inspection data to improve flaw detection and characterization— data that will support more certain failure pressure calculations for inspected pipelines
-  Get NDA(s) in place with ILI service provider(s) who can share data on pipelines located in the U.S., or representative of those in the U.S.
-  Transfer data from operator partner(s) to the PNNL Data Stewardship Board to “de-identify” the data, if required

FY 2021 - FY 2022

Expected Key Outcomes:

- 1) New data-driven algorithms produced by applying machine learning to past NGT pipeline ILI signal data to reveal relationships between signal characteristics and physical flaw profile

Primary Tasks:

-  Apply ML to pipeline signal data to produce new data-driven flaw detection and “reconstruction” models

FY 2022

Expected Key Outcomes:

- 1) Verified and validated data-driven model of flaw detection and reconstruction
- 2) Ancillary outcome: specific data that could be collected in the future, using existing or new or additional ILI sensors

Primary Tasks:

- *Verify* new data-driven models using ~10% of training set data
- *Validate* new data-driven models using another ~10% of reserved data sets to determine if model predictions reflect ground-truth answers with an acceptable level of accuracy, e.g., 90+%

FY 2023

Expected Key Outcome:

**Phase I Milestone:** Produce Alpha Data-driven Model for ILI Signature Screening (Flaw Detection) & Flaw Reconstruction v0.1 (research-grade prototype software intended for first round of alpha testing)

Primary Tasks:

- Test the model (via hindcasting) to determine accuracy of predictions
- Refine, re-test until alpha ready



# Prognostic Model

## Flaw initiation time, growth rate, & significant corrosion factors









- Apply machine learning to:
  - large set of flaw characteristics in existing ILI reports for a diverse set of pipelines over their service lives
  - corresponding pipeline operating history, environmental conditions, pipeline material properties and construction information
    - known/suspected factors that affect initiation and progression of corrosion
    - based on real-world experience and knowledge (domain expertise)
- Once tested and validated, the model can be used to predict corrosion initiation/growth rate for a pipeline to calculate remaining useful life
  - based on an inspected pipeline's current condition, operating history, and other pipeline attributes – obvious and non-obvious and significance of each – to be uncovered during the project
  - may be possible to infer remaining life of *uninspected* pipelines based on attributes shared with inspected pipelines

### Status:

- Received ILI report data and corresponding pipeline attribute data (May-July 2021)
- Curation of data to prepare for prognostic model training is in-progress
- **Seeking historical soil/environmental and land-use data to add to training data**

# Prognostic Model Progress

## Phase I (current project scope): *Development Phase Between PNNL and Pipeline Operator Partner(s)*

<p>FY 2019 - FY 2021</p> <p><u>Expected Key Outcomes:</u></p> <ol style="list-style-type: none"> <li>1)  <b>Framework for machine learning</b></li> <li>2)  <b>Strategic partnership(s) with NGT pipeline operator(s)</b></li> <li>3)  <b>Sharing of NGT pipeline data sets for a representative sampling of NGT pipelines located across the U.S. needed to train robust machine learning algorithms</b></li> </ol>	<p>FY 2021 - FY 2022</p> <p><u>Expected Key Outcomes:</u></p> <ol style="list-style-type: none"> <li>1) New data-driven algorithms produced by applying machine learning to past data to reveal corrosion evolution rates, initiation times, failure pressure/age and their correlation with key properties of pipelines (essential variables)</li> <li>2) Down-selected physics-based models of corrosion to hybridize with the new data-driven model of corrosion</li> </ol>	<p>FY 2022</p> <p><u>Expected Key Outcomes:</u></p> <ol style="list-style-type: none"> <li>1) Verified and validated data-driven model of corrosion evolution</li> <li>2) Ancillary outcome: specific data that could be collected in the future, using existing or new process monitoring or ILL sensors</li> </ol>	<p>FY 2023</p> <p><u>Expected Key Outcome:</u></p> <p><b>Phase I Milestone:</b> Produce Alpha Hybrid Model for Corrosion v0.1 (research-grade prototype software intended for first round of alpha testing)</p>
<p><u>Primary Tasks:</u></p> <ul style="list-style-type: none"> <li>•  Build framework for hybrid data-driven, physics-based model, starting with corrosion</li> <li>•  Develop partnerships with NGT pipeline operator(s) who have interest in applying machine learning (ML) to historical pipeline data to improve certainty of time-to-failure (TTF) projections (prognostics) for inspected and un-inspected pipelines</li> <li>•  Get NDA(s) in place with operator partner(s) who can share data on pipelines located in the U.S., or representative of those in the U.S.</li> <li>•  Transfer data from operator partner(s) to the PNNL Data Stewardship Board to “de-identify” the data, if required</li> </ul>	<p><u>Primary Tasks:</u></p> <ul style="list-style-type: none"> <li>•  Apply ML to pipeline data sets to in-line inspection report data, etc. to produce new data-driven corrosion evolution rate models, data-driven corrosion initiation time models, and data-driven failure pressure/age models</li> <li>• Perform “hindcasting” with pre-existing corrosion evolution rate models to determine which ones yield the most accurate results and should be considered for “hybridization” with the new data-driven corrosion evolution rate model</li> </ul>	<p><u>Primary Tasks:</u></p> <ul style="list-style-type: none"> <li>• <i>Verify</i> new data-driven models using ~10% of training set data</li> <li>• <i>Validate</i> new data-driven models using another ~10% of reserved data sets to determine if model predictions reflect ground-truth answers with an acceptable level of accuracy, e.g., 90+%</li> </ul>	<p><u>Primary Tasks:</u></p> <ul style="list-style-type: none"> <li>• Hybridize validated data-driven corrosion models with down selected pre-existing physics-based corrosion evolution models</li> <li>• Test the hybridized models (via hindcasting) to determine accuracy of predictions as compared with individual data-driven and physics-based models</li> <li>• Add statistical models</li> <li>• Test the statistical models (via hindcasting)</li> <li>• Refine, re-test until alpha ready</li> </ul>

# Envisioned Ultimate Use

Technical Status  
Accomplishments  
Lessons Learned  
Synergy Opportunities  
Project Summary  
Appendix

Step 1: Obtain PDFs from Hybrid Model  
Operators use the PHD software to obtain PDFs for different pipeline conditions (e.g., corrosion, stress, etc.), as in service pipelines, customized for each pipeline segment.

## Research (Us)

Inspected  
Pipelines

- PDF for corrosion evolution rate
- PDF for corrosion initiation time
- PDF for baseline failure pressure/age

### Pipeline Health Display V1.0 (Commercial grade executable)

**Inputs:** Essential Pipeline Variables

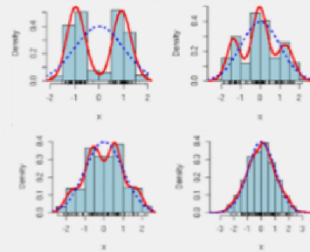
Material Parameter

Soil	Temp's Modulus	0.0e+000 Pa
Friction's Ratio	0.0	
Density	1900 kg/m <sup>3</sup>	
cohesion's Value	18000 Pa	
Angle of Friction	0.471 rad	
Pipeline	Temp's Modulus	0.0e+000 Pa
Friction's Ratio	0.0	
Density	7800 kg/m <sup>3</sup>	
cohesion's Value	1.6e+008 Pa	
Temp's Modulus	0.0e+000 Pa	

Working Condition of Pipeline

Diameter	1.000 m	Thickness	0.005 m
Depth	0.0 m	Inner Pressure	1e+007 Pa

**Outputs:** Probability density functions for corrosion evolution rate; corrosion initiation time; baseline lifecycle or failure age/pressure; and condition

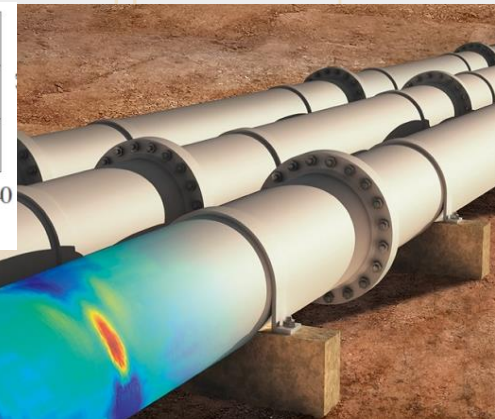
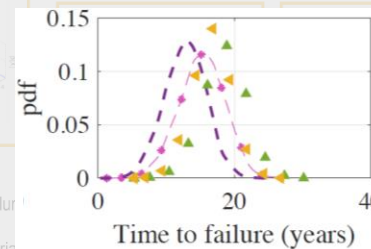
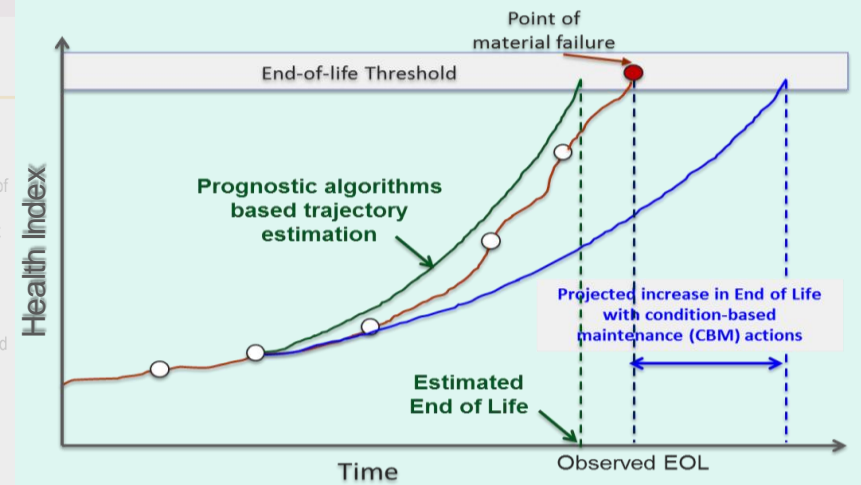


Un-Inspected  
Pipelines

- PDF for corrosion evolution rate
- PDF for corrosion initiation time
- PDF for pipeline condition
- PDF for baseline failure pressure/age

Step 3: Calculate Pf, POF & Risk  
Use flow characteristics to determine Pf and use Pf to calculate POF for different SysOPs

## Operations (Industry)



\*Pf calculations are updated as appropriate, i.e., after ILI or when new data-driven Pf PDFs become available (following new version releases).

Amaya-Gómez R, M Sánchez-Silva, E Bastidas-Arteaga, F Schoefs, and F Munoz. 2019. "Reliability assessments of corroded pipelines based on internal pressure – A review." *Engineering Failure Analysis*.

# Accomplishments To-date

## Summary:

### 1. Data Received from Industry Partners

- **One Major ILI Service Provider / MFL Tool Developer**

- NDA in effect since July 2020
- Received multiple MFL signal datasets for engineered & real pipeline corrosion (400 MB)
- Working collaboratively with ILI company to build physics-based, data-driven model to detect and digitally reconstruct flaws

- **Two Major Natural Gas Transmission Pipeline Operators**

- NDAs in effect since March 2021
- Received > 2 GB data so far
- Working collaboratively to build physics-based, data-driven model of corrosion initiation time and growth rate

### 2. Began training diagnostic model with ILI signal data

### 3. Began curating training data set for prognostic model

#### Partnership at a Glance

NDA/MTA signed and in effect

**2** Major Pipeline Operators

**1** Major In-line Inspection Service Provider

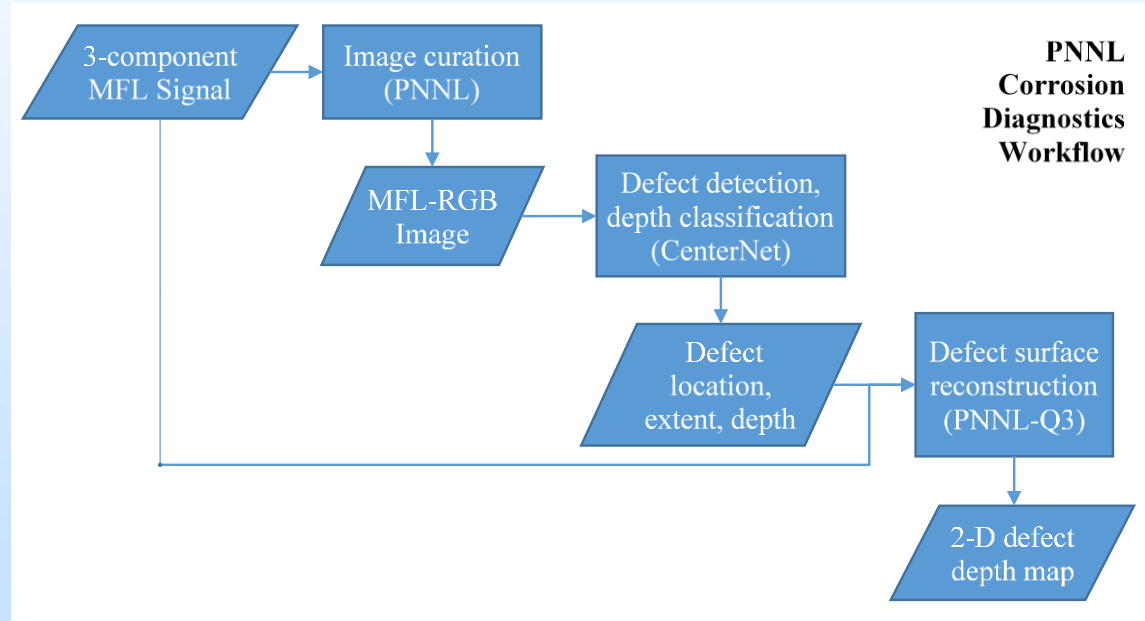
**1** Key Industry Research Consortium

# Diagnostic Model

## Diagnostic Corrosion Analysis Pipeline

### Motivations:

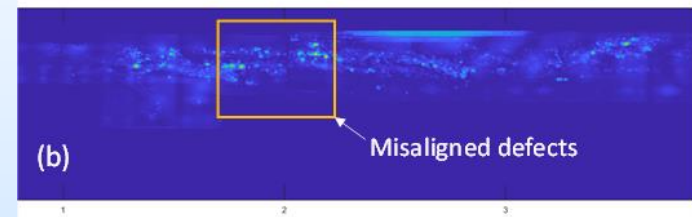
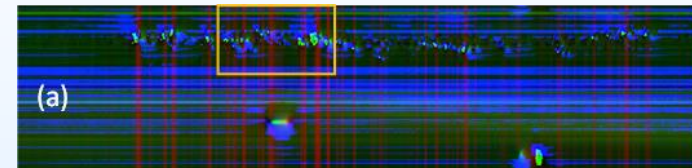
- A key task in pipeline lifecycle analysis is defect alignment over time.
- Arbitrarily picking defects is prohibitively time consuming and subjective.
- More intensive surface analysis requires a more focused region
- More accurate failure pressure calculations



# Diagnostic Model

## Real Corrosion Defect Data Curation

- Unlike engineered defects, real corrosion defects are usually not distinct however their exact size and location are still needed for training the models.
- Aligning the MFL scan data with ground truth is a challenge and vendors use proprietary software.
- Alignment based on partner provided information resulted in offsets (discontinuities), so the team developed landmark (feature) correspondence-based technique to align the MFL scans with laser ground truth data .
  - This will not only help in identifying appropriate depths and bounding boxes for training the ML models but also assist in the defect reconstruction from MFL data which is one of the partner's key interests.



(a) MFL Scan (b) Aligned ground truth (laser) image with partner data



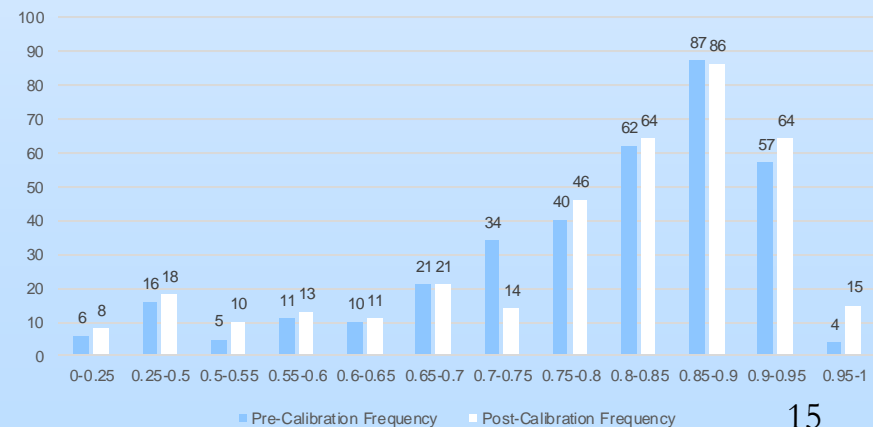
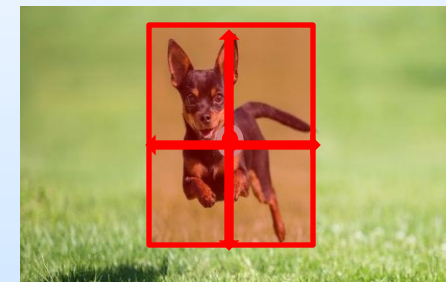
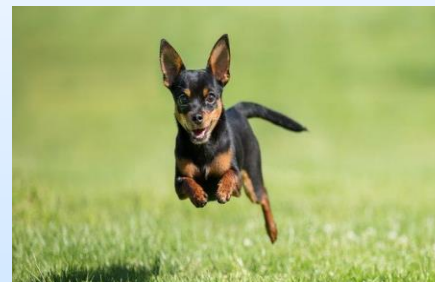
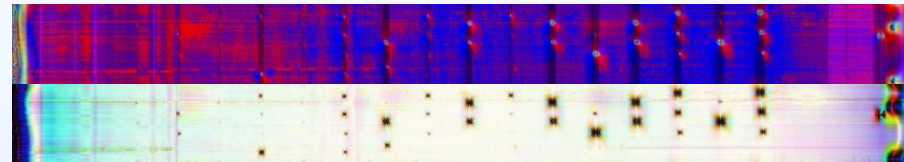
(c) Laser scan data (d) MFL Scan with matching features and misaligned positions



# Diagnostic Model

## Defect Detection and Depth Classification

- 3-component MFL signals from the ILI data provider rescaled to RGB color images (a)
- CenterNet image identification module (b) trained on human-curated defect extent and depth data (MFL) to
  - Predict tight bounding box around defect
  - Estimate defect depth (binned every 10%)
- Defect boundary identification accuracy of 81% over defect penetration depths evenly distributed from 10% to 100%
  - Increased from 63% before expanding dataset and refining RGB encoding.
- Developed RGB defect feature amplification technique which increased the detection percentage by an additional 5% (18)

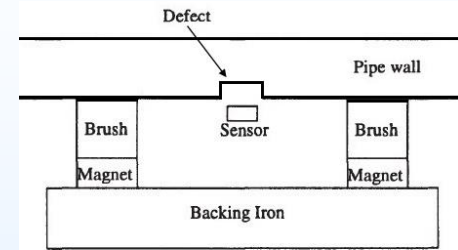


# Diagnostic Model

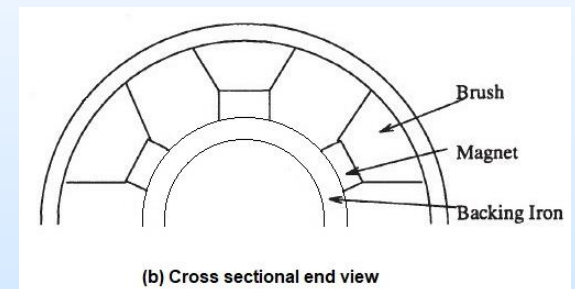
## Validated FE model for supplemental training data

- There is limited field MFL data that correlates measured MFL signals to known defect profiles. Finite element modeling (FE) can simulate MFL inspections and generate synthetic but realistic testing data for various defect geometries and dimensions. This will provide more training data for ML.
- This task will set up finite element models for the ILI MFL PIGs used in the field.
- Currently working with ILI partner on model parameters and validation against real defects.
- Synthetic data supplements/interpolates an array of real training data over range of defect profiles, pipe sizes, and inspection speeds.

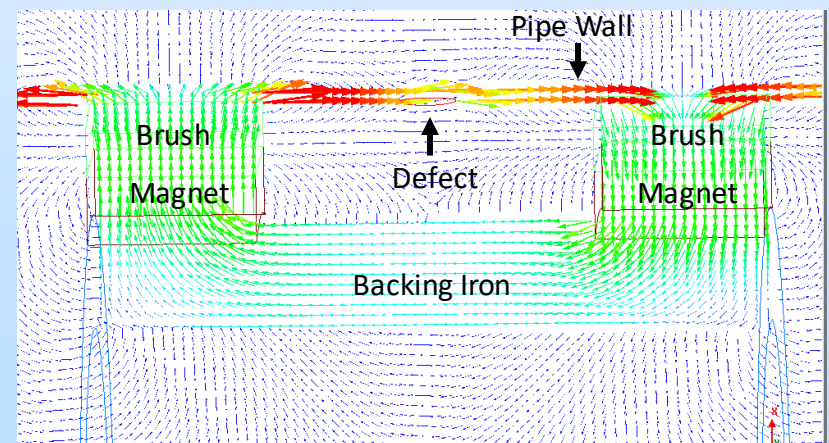
MFL PIG Setup



(a) Cross sectional side view



(b) Cross sectional end view



Simulated Magnetic Flux Density Using ANSYS Maxwell



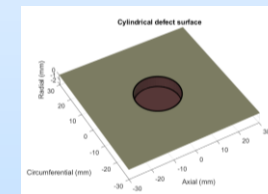
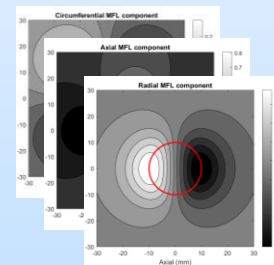
# Diagnostic Model

## Machine Learning for Defect Surface Reconstruction

- Reconstruct corrosion defect surface features from diagnostic inspection to assess present state of pipeline health
- Compiled and analyzed 5 diagnostic models spanning several algorithms and fidelities
- State-of-the-art (RBF) ML techniques identified in most moderate-high performance approaches
- Now evaluating on real/engineered datasets for final selection and optimization

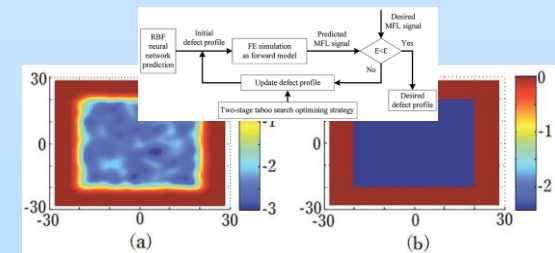
### Model identification and selection grid

Model	Type	Fidelity	Cost
Dutta (2009)	Physics	Low	Low
Hwang (2000)	ML – RBF	Med	Low
Chen (2014)	ML – RBF	Med	Med
Han (2017)	ML + Physics	Med	High
Chen (2016)	ML + Physics	High	High



Dutta (2009):  
Physics model  
(simplified)

Chen (2016):  
Iterative ML+Physics  
hybrid



# Prognostic Model

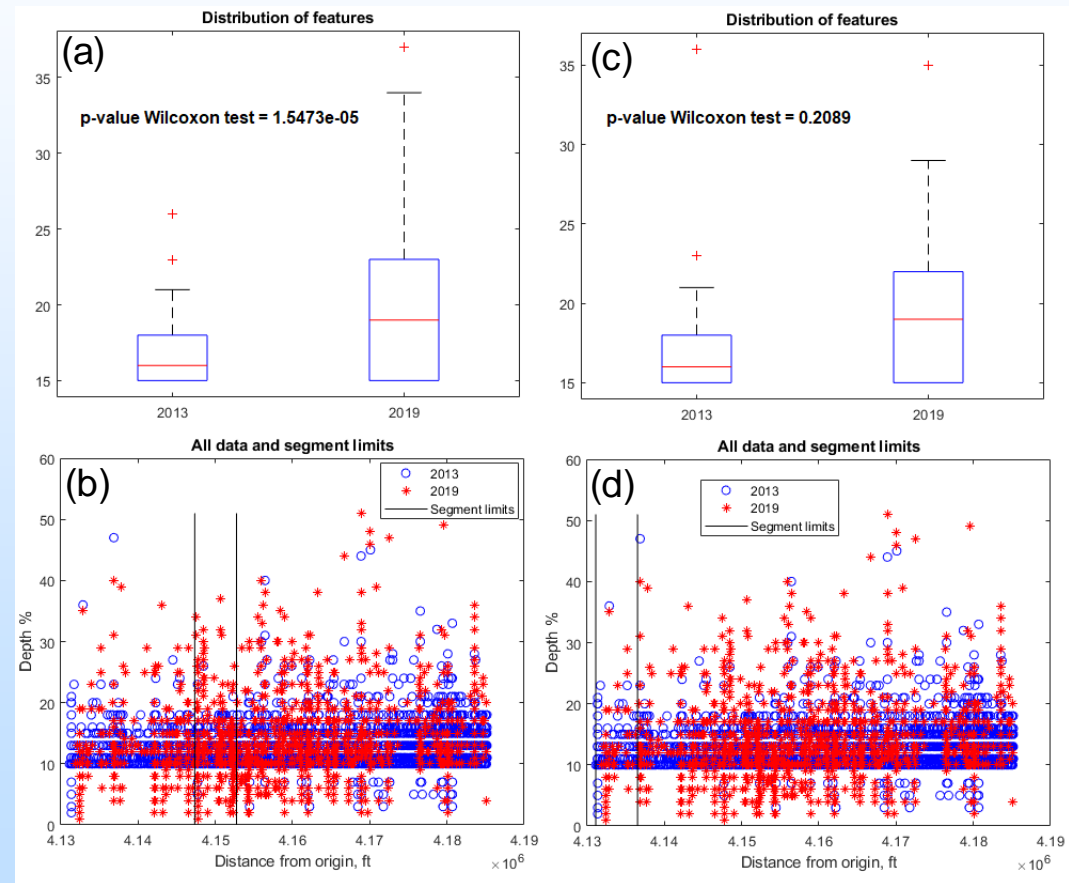
## Prognostics: Data Sets and Curation

- Successfully collaborated with two prominent O&G operators and received periodic ILI data and associated metadata in various database formats.
- The analysis of data indicated growth rate estimates based on as received data is not a straightforward process and significant curation is needed to make the data conducive to ML applications.
  - Both defect-to-defect matching and segment wise aggregate methods are considered.
  - Detection thresholds and defect matching introduce significant uncertainty.
- Establishing secured, centralized databases with consistent format and relevant attributes where data extracted from various data sets will be merged for data driven model development.
- Efforts are also underway to supplement the ILI and metadata with other geospatial corrosion driving factors such as soil physical and chemical compositions, land use type, proximity to AC inference, environmental factors
- Models compatible for various scenarios also being investigated so that prognostic models based on single or periodic inspections could be developed if necessary.

# Prognostic Model

## Segment Aggregate Based Comparison

- Advantage: No defect matching
- Wilcoxon test
  - tests the hypothesis that the median depth % metal loss is equal in 2013 and 2019
  - p-values < 0.05 indicate hypothesis should be rejected
  - Any pipe segment can be analyzed to find significant differences
- Pipe segments from different locations of the same line indicated significant and not-significant differences in median depths when defects are aggregated.
  - This approach could assist in identifying locations of significance from metals loss perspective.

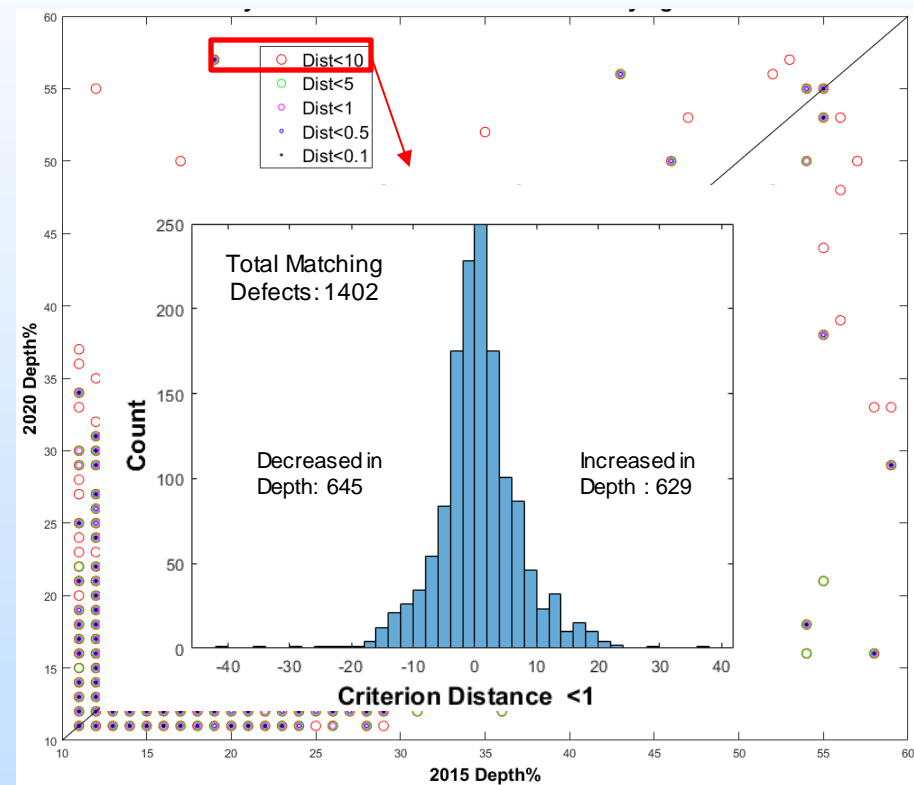


Segments with significant (a) & (b) and not-significant (c) & (d) aggregated defect depths from periodic scans of a pipeline

# Prognostic Model

## Defect Matching and Impact on Corrosion Rate Estimates

- Matching of defects is one of the most challenging aspects in data curation
  - Different tools with different tolerances and biases were used
  - Uncertainties in defect characterization could be larger than growth over periodic scans
  - Different formats could be followed by ILI service vendors in reporting ILI data.
- Data from matched defects indicated defects both increasing and decreasing in depth which may introduce significant uncertainty in growth rate estimates
  - Better matching and filtering criteria under development with feedback.
  - Aggregate depth methods are also being investigated.



Defect depth comparison from 2015 and 2020 ILI data based on a matching distance.

No dominant growth or shrinkage patterns observed.

# Lessons Learned

- Research gaps/challenges
  - Obtaining accurate ground-truth from industry data
  - Flaw matching/correlation for a given pipeline between inspection intervals
- Unanticipated research difficulties
  - Extra time to build trust/credibility with industry partners, receive data
    - Took a year longer than planned to receive data for diagnostic model
    - Took 1.5 years longer than planned to receive data for prognostic model
    - Industry open to partnering on this project; some have attempted ML themselves or with others and are still open/optimistic about continuing
- Technical disappointments
  - Managing expectations of maturity level of models at the end of Phase I
    - Appropriate for “alpha testing,” not complete technology transfer to industry
- Changes that should be made next time
  - Create three versions of data requirements documents
    - For initial “socialization” of concept with potential partners: high-level visual summary of data types & variety/quantity of each data type (pipeline dia., mileage/length, #inspection intervals)
    - For internal project use: add'l details and specs. on allowable/acceptable errors (e.g. position)
    - For communications with actual partners: same s, file format(s), variety/diversity of pipe size, quantity of replicates (ILI)

# Synergy Opportunities

- Synergy with other Data Science/Machine Learning projects
  - Seeking historical soil/environmental and land use data to add to prognostic model training data
    - Could leverage previously curated data from multiple public/private databases (e.g., SSURGO)
- Synergy with other NGT pipeline integrity projects
  - Complementary pipeline flaw aggregation/matching techniques
  - Identifying new discoveries and insights about pipeline materials to inform corrosion modeling

# Project Summary

- Highlights (Key Findings):
  - Established partnerships with 3 O&G partners & received data for ML model training
  - Developed end-to-end diagnostics capability for MFL ILI data:
    - Defect detection, curation, simulation, and characterization workflows/modules
  - Paved way for prognostics analysis:
    - Secure partner-specific databases, geospatial data sources, analytics for time-dependent statistical defect growth
- Next Steps:
  - Refine defect diagnostics accuracy:
    - Augmented real/simulated corrosion training dataset, improved hybrid modeling
  - Prognostics pipeline development:
    - Time-dependent geospatial defect database, prognostics model selection/implementation

# Appendix

---

- Benefits to Program
- Project Overview
  - Summary, Technical Objectives, Tasks in SOPO
  - Program Alignment
  - Success Criteria
- Organization Chart
  - Project team, Organization, and Participants
- Gantt Chart
- Bibliography



# Benefit to the Program

Leverage advances in machine learning and predictive analytics to advance the state of the art in pipeline infrastructure integrity management using forecasted (predicted) pipeline condition, using large sets of pipeline integrity data and continuous operational data generated by oil and gas (O&G) transmission pipeline operators.

Goals: Develop hybrid physics-based/data-driven models that help NGT pipeline operators make integrity management decisions that ultimately:

- Reduce pipeline incidents/failures that compromise the overall reliability of the “gas grid,” result in supply interruptions, and compromise safety

# Project Overview

## Summary

---

- Project Objective:
  - Develop research-grade data-driven/physics-based diagnostic/prognostic models
- Project Scope:
  - Gather past data to facilitate machine learning (initial focus: natural gas transmission, external corrosion)
  - Develop, verify and validate the research-grade diagnostic/prognostic models
- Anticipated Project Outcomes:
  - Reports on model performance and “research-grade” prototype models (software) ready for alpha ( $\alpha$ ) testing in Phase II
    - Example of  $\alpha$  testing: make predictions on pipelines that are already slated to be removed from service so true condition can be compared with predicted condition
  - Recommendations for what/how data should be collected to improve prediction abilities – i.e., what measurements next-generation “smart pipes” should perform

# Project Overview

## Technical Objectives

---

- 1. Team with NGT pipeline industry and apply ML to historical NGT pipeline data sets**
  - ILI data (ILI tool signal data or flaw sizes listed in ILI reports)
  - Pipeline attributes (material, environmental conditions, construction history, etc.)
- 2. Uncover “novel signatures” in data sets to gain new insights on current & future pipeline condition**
  - Non-obvious ILI signal features used to increase flaw detection probability, resolution & accuracy of flaw size (MFL or UT)
  - Non-obvious relationships between pipeline corrosion initiation time, corrosion rate, and pipeline properties/attributes
- 3. Use novel signatures to build model**
  - Hybrid physics-based, data-driven diagnostic model for assessing current pipeline condition
  - Hybrid physics-based, data-driven prognostic model for predicting future pipeline condition
- 4. Generate algorithms with models (to prepare for alpha tests)**
  - Outcomes will determine if follow-on phase (alpha testing, improvements, beta testing) is warranted

# Project Overview

## Tasks in Statement of Project Objectives

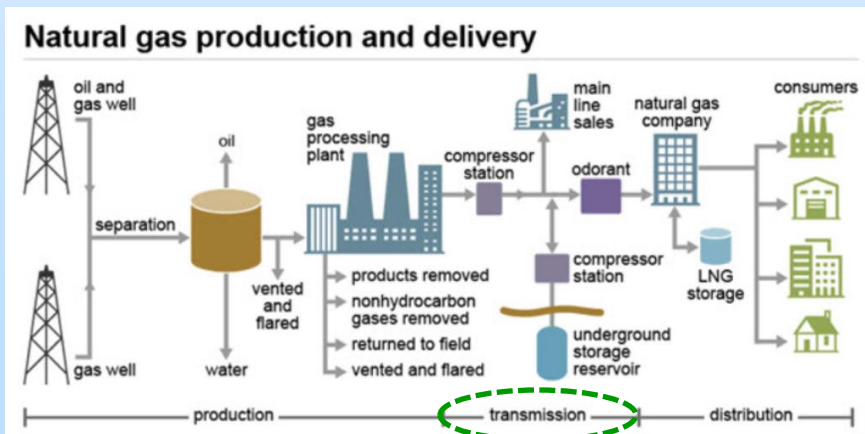
---

- Task 1: Data Requirements and Accessibility
  - Develop data requirements
  - Outreach for data set availability and accessibility
  - Data set selection
- Task 2: Diagnostic Algorithm Development
  - ML based algorithms for signature extraction
  - Algorithms to use these signatures for diagnostic purposes (detection and characterization of degradation)
- Task 3: Prognostic Framework Algorithm Development
  - ML based algorithms to predict remaining service life
- Task 4: Pipeline Reliability & Lifecycle Health Management System
  - Integrate algorithms from Tasks 2 & 3
- Task 5: Sensors SME Team
  - Expert Panel
  - Sensor Technology Database

# Project Overview

## Program Alignment

- DOE-FE Program Alignment: Natural Gas Infrastructure
  - Natural Gas Technologies (NGT) R&D
  - The NGT R&D program is aligned with the President's objectives to strengthen natural gas pipeline reliability and ensure infrastructure security, and research advanced materials and sensor technology to address natural gas infrastructure reliability, public safety, and operational efficiency
    - <https://www.energy.gov/fe/natural-gas-technologies-rd>
    - <https://netl.doe.gov/oil-gas/ngi>



# Project Overview

## Success Criteria

1. **Early go/no-go point:** Partner with NGT industry to obtain data sets to develop the data-driven aspects of the diagnostic and prognostic models
  - ✓ **Satisfied**
  - One Major ILI Service Provider / MFL Tool Developer
  - Two Major Natural Gas Transmission Pipeline Operators
  - Working collaborative with all three to develop the models
2. **Target goals for the diagnostic and prognostic models:**
  - a) a reduction in [external corrosion] flaw detection false alarm probabilities to less than 1% (for MFL-detectable flaw sizes of  $\geq 20\%$ )
  - b) maximum predictive time to critical flaw size uncertainty of 5% over an inspection cycle
  - c) an [external corrosion] flaw detection and localization accuracy exceeding 90% (at a 95% confidence level) for high consequence regions of the pipeline network

# Organization Chart



Kayte  
Denslow, PI  
NDE,  
Sensors



Steven  
Rosenthal,  
Task Lead  
(Math/ML)



Naveen  
Karri, Task  
Lead (Mech  
Eng.)



Arun  
Veeramany,  
Risk &  
Reliability,  
ML



Alejandro  
Heredia-  
Langner,  
Statistics,  
ML



Xinming  
Lin,  
Data  
Science



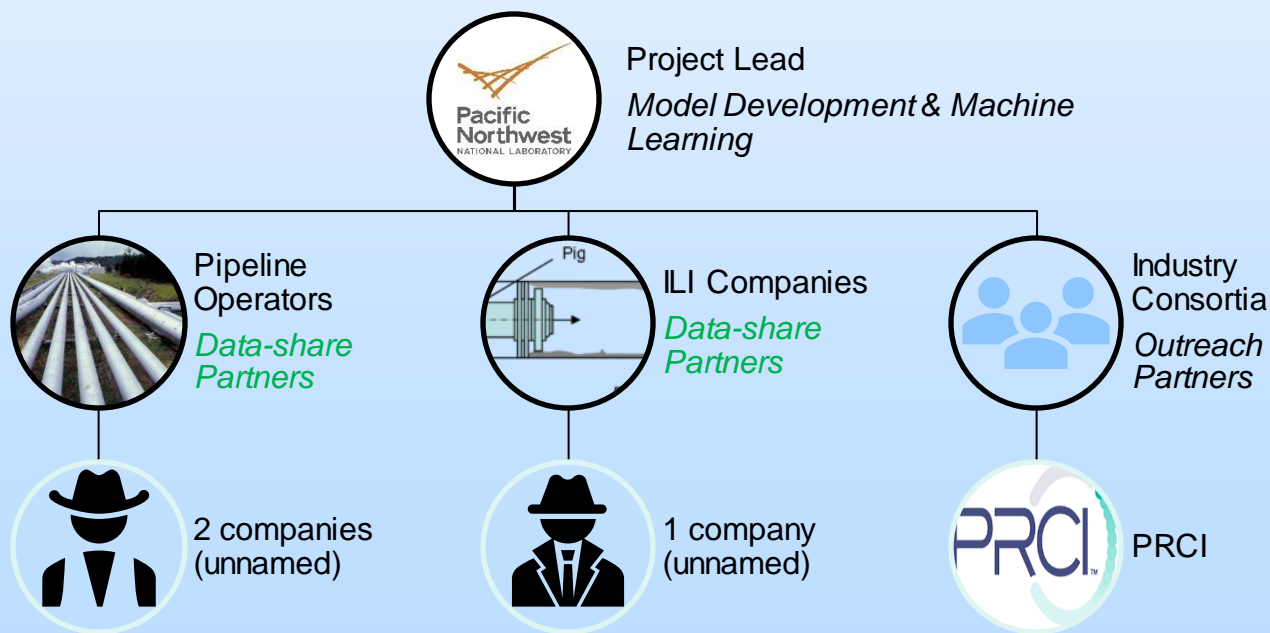
Juan  
Brandi-  
Lozano  
Math/ML/  
Industry



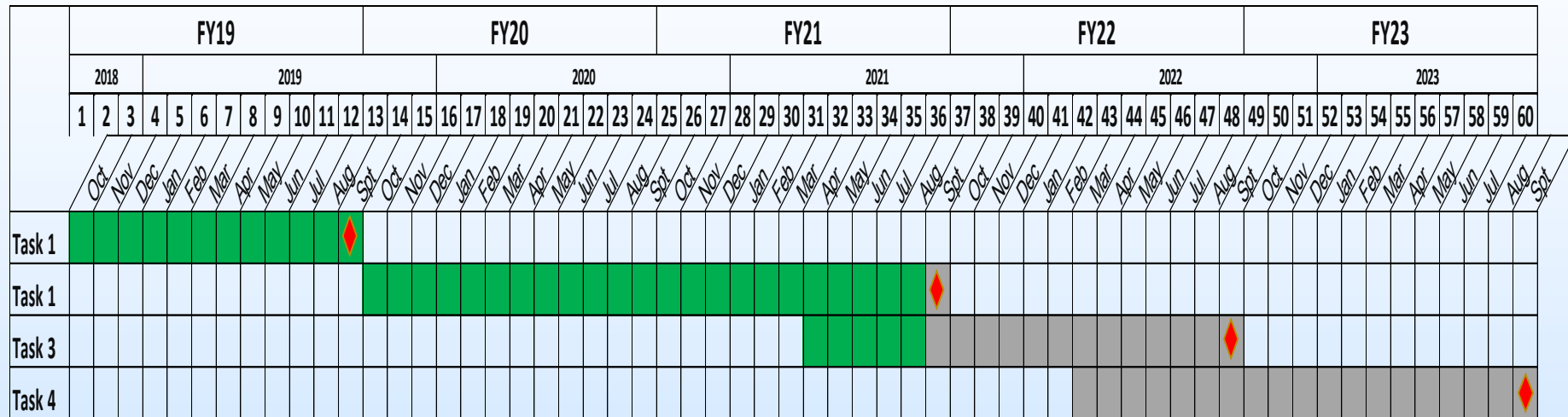
Yanming  
Guo,  
NDE&T,  
Ultrasonics,  
MFL



Angela  
Dalton,  
PM

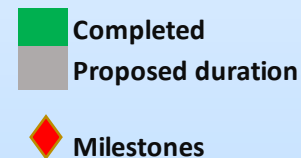


# Gantt Chart



## List of Deliverables:

- Deliverable 1: Industry Data Workshop Summary
- Deliverable 2: Diagnostic Algorithm Evaluation Report
- Deliverable 3: Prognostic Algorithm Evaluation Report
- Deliverable 4: Pipeline Reliability and Lifecycle Health Management System Design Report





# Bibliography

---

- Presentations at REX2020, PRCI 2020 Webinar
- Training of models is in-progress, so peer-reviewed publications will be produced in the future