**CAAP 1st Annual Report**

Date of Report: *10/14/2022*

Prepared for: *U.S. DOT Pipeline and Hazardous Materials Safety Administration*

Contract Number: *693JK32150001CAAP*

Project Title: *Pipeline Risk Management Using Artificial Intelligence-Enabled Modeling and Decision Making*

Prepared by: *Bingyan Cui (PhD student), Emad Farahani (PhD student), Jingnan Zhao (Postdoc), Dr. Hao Wang (PI), Dr. Qindan Huang (Co-PI)*

Contact Information: Hao Wang, hw261@soe.rutgers.edu, 848-445-2874

Qindan Huang, qindan.huang@marquette.edu, 414-288-6670

For quarterly period ending: 10/1/2021 - 9/30/2022

**Business and Activity Section**

# Contract Activity

* Contract Modification

N/A

* Student Mentoring

Two PhD students are mainly working on the project tasks. The PhD student at Rutgers University, Bingyan Cui, led the work on literature review, data collection and analysis, and development of machine learning models for defect growth prediction. The PhD student at Marquette University, Emad Farahani,led the work on development of Bayesian statistics model for defect growth prediction.

The postdoc at Rutgers University, Jingnan Zhao, helped on machine learning model development.

* Educational Activities

The PI introduced the knowledge of pipeline integrity management system in the graduate course – *Infrastructure Management System* taught at Rutgers University.

The Co-PI introduced the knowledge of pipeline failure prediction and risk management in the graduate course - *Engineering Risk Analysis* at Marquette University.

* Outreach Activities

The research team collected pipeline in-line inspection (ILI) data from the industry partners for developing defect growth models.

# Financial Summary

* Federal Cost Activities

The PI has spent time working on the project, but it was charged through cost-sharing not from this project during this reporting period. The Co-PI has started working on the project and charged her time through cost-sharing. The graduate students have spent time on the project, but only partial amounts of the salary were charged from the project since the graduate students were also funded by the fellowship from university or external sources.

Cost shear activities were provided by the PI as specified in the proposal.

# Project Schedule Update

The schedule of research tasks is shown in the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| **Tasks** | **Year 1** | **Year 2** | **Year 3** |
| ***Task 1*** *Literature Review*  |  |  |  |  |  |  |  |  |  |  |  |  |
| ***Task 2*** *Data Collection from Literature and Industry Partners* |  |  |  |  |  |  |  |  |  |  |  |  |
| ***Task 3*** *Data-Driven Probabilistic Modeling of Pipeline Defects* |  |  |  |  |  |  |  |  |  |  |  |  |
| ***Task 4*** *Quantification of Probability of Failure* |  |  |  |  |  |  |  |  |  |  |  |  |
| ***Task 5*** *Decision Making with Reinforcement Learning* |  |  |  |  |  |  |  |  |  |  |  |  |
| ***Task 6*** *Final Report and Presentation* |  |  |  |  |  |  |  |  |  |  |  |  |

# Status Update of the 4th Quarter Technical Activities

In the 4th quarter, the research team proposed a power-law model formulation as a function of time to predict the nonconstant growth of pipe defects (maximum depth and length of corrosion). Bayesian statistics was used to estimate model parameters. The technical details of the work are provided in 4th quarterly report.

**Detailed Technical Results in the Report Period**

# (a) Background and Objectives in the Annual Report Period

The objective of 1st year work is to collect ILI inspection data from pipeline operators, conduct statistical analysis of pipeline defects, and started developing probabilistic model of defect growth.

# (b) Research Progress

The following provides the summary of work for each task. The detailed data and analysis can be found in the quarterly reports.

*Task 1 Literature Review*

Literature review has been conducted for the prediction models of pipeline defect growth. First, traditional prediction models such as Markov models, inverse Gaussian process-based models, and geometric Brownian motion models with drift were discussed. Then, artificial intelligence-based models, including ANN and BNN models, were highlighted. Finally, the advantages and limitations of these models were discussed.

*Task 2 Data Collection from Industry Partners*

Field inspection data were collected from two industry partners for data analysis.

1) The first set of data is for a 12-mile steel transmission pipeline that was originally constructed in 1974 in US. The pipe outside diameter is 20-30 inches and the wall thickness varies from 0.375 to 0.562 inch. The pipe materials are mainly 5L×42 with some replacement segments being 5L×60. The in-line inspection (ILI) were conducted using Magnetic Flux Leakage (MFL) inspection tools in 2005, 2012 and 2016. The MFL tool provides measurements of wall thickness and metal loss depth/length/orientation. The main defect is external metal loss due to corrosion. Internal metal loss is mainly caused by manufacturing defects and the variations of wall thickness for seamless pipe, which should not be considered as internal corrosion. Crack is very rare since the MAOP is maintained smaller than 40% of design pressure. A few dent defects were observed.

* The external corrosion defects were analyzed using k-means clustering based on metal loss depth and length. The results show that the corrosion defects can be divided into four clusters along the length of pipeline, which indicate the potential impact of soil environment on external corrosion. However, the soil survey data were not available.
* Although the number of defects increased from 2005 to 2016, the average corrosion depth did not increase. This was because the generation of large numbers of small defects in 2012 and 2016.
* The relative distance to the girth weld number was used to locate the defect location and only the data that shows the continuous growth trend of maximum corrosion depth over inspection years were selected for further analysis. The maximum corrosion depths in the subset were ﬁtted to the Gumbel distribution with location and scale parameters, which shows an increasing trend over inspection years. In addition, the axial and circumferential locations of defects are analyzed through density plots.

2) The second set of data is for a buried steel (grade X52) pipeline located in the southeast of the Gulf of Mexico. This studied pipeline with a total length of 112 km has been in service since 1969. It has an outside diameter of 457.2 mm and wall thickness of 6.4mm. The pipe was inspected in 2005 using MLF technology and in 2010 using straight beam ultrasound technology. Along the pipeline, soil properties were measured at every 200m. Soil properties include Eh (soil redox potential), resistance, resistivity, pH, CO32- concentration, HCO3- concentration, Cl- concentration, SO42- concentration, soil moisture, and soil type. In ILI inspection, external corrosion information was collected, including metal loss depth, defect length, and orientation. The pipeline coating condition data were not available.

* The metal losses (external corrosion depth) show a general trend of increase between two inspection runs.
* Although it is difficult to find the exactly matched defects at the same location, it is evident that the median and mean of metal loss from the second inspection are greater than those from the first ILI run. This reflects corrosion behavior caused by degrading surfaces due to soil environment. However, this trend was not obvious for defect length, which indicate different development trends of defect depth and length due to external corrosion.
* The electrical and chemical properties of soil were found having large variations. There was one soil zone shows very high Cl- concentration and the calculate corrosion rate was high. However, simple correlations between individual soil parameters and corrosion rates were not found.

It is expected that the corrosion depth would increase over years if no repair is placed. However, this trend was not always observed at each inspection location for both pipelines. The variations can be caused by 1) the changes in measurement accuracy, tolerance, and reporting criteria of ILI tools and the advances in technology in different inspections, and 2) the mismatch of absolute distance from odometer readings during different inspections. The research team will further look into the collected ILI data with industry partners. The dataset used in the further analysis will be prepared based on model assumptions and needs.

Due to the availability of data collected from industry partners, the proposed analysis in this project will focus on predicting defect growth due to external corrosion considering the effects of soil properties mainly. It is recognized that the accuracy of prediction model may be affected by the number of affecting factors included in the model. If the model is used for other pipelines, it can be modified to consider other factors affecting external corrosion behavior, such as pipe material, manufacturing process, coating, catholic protection (CP), and operating conditions, if the information is available.

*Task 3 Data-Driven Probabilistic Modeling of Pipeline Defect Generation and Growth*

The second set of data from Gulf of Mexico was used to analyze defect generation and growth models for both corrosion maximum depth and length, a power-law function of time model formulation, as expressed in Equation (1), is adopted which considers nonconstant damage growth rate over time. A Poisson process is considered for the occurrence of defects, indicating the initiation time of each individual defect is considered to follow a Gamma distribution. Bayesian statistics is used to assess the joint probability density function (PDF) of the unknown model parameters and the constructed likelihood function considers the correlation between defect depth and length growth. The predicted corrosion depth and length and the associated actual measured quantities obtained through ILI are compared.

# (c) Future Work

Future work will be conducted in the following aspects to further develop probabilistic models of pipeline defects:

1. The soil parameters will be incorporated into the probabilistic growth models to consider the variation of defects along the pipeline. In addition, the appropriate transformation of defect quantity (depth and length) to ensure the assumption satisfaction of Normal model error distribution.
2. The Bayesian Neural Network (BNN) based algorithm will be developed to predict defect growth considering soil parameters as input features. Generative Adversarial Networks (GANs), a deep-learning-based generative model will be explored to enhance the size of dataset.