**CAAP Quarterly Report (updated)**

**1/24/2022**

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

*Contract Number: 693JK32150001CAAP*

*Prime University: Rutgers University*

*Prepared By: Dr. Hao Wang,* *hwang.cee@rutgers.edu**, 848-445-2874*

*Reporting Period: 9/30/2021 – 12/31/2021*

**Project Activities for Reporting Period:**

*Kick-off Meeting*

The project kick-off meeting of this project was conducted on Oct. 29, 2021 remotely. The attendees include Zhongquan Zhang (PHMSA), Nathan Schoenkin (PHMSA), Dr. Hao Wang, (PI), Dr. Qindan Huang (Co-PI), Binfyan Cui (PhD student at Rutgers), Jingnan Zhao (Postdoc at Rutgers), and Kiswendsida Jules Kere (PhD student at Marquette). The PI first gave a presentation on the overall project objective, research tasks, and student mentoring plan. After that, the research team had some discussions with PHMSA representatives.

Based on the discussions in the meeting, two action items were conducted by the research team.

* Task 3.2: Describe the difference of task 3.2 as compared to the existing PHMSA Core project (Mapping Indication Severity Using Bayesian Machine Learning from Indirect Inspection Data into Corrosion Severity for Decision-Making in Pipeline Maintenance) on model development for indirect measurement data. Zhongquan suggested to review the existing PHMSA core project which will be completed in 2022.

The research team reviewed the information (public abstract and quarterly reports) of existing PHMSA Core project that is available online in the PHMSA Research & Development Program. The datasets were linked with the nearby soil survey data, large scale vegetation and precipitation data, indirect inspection data, and in-line inspection data. The indirect inspection data include Close Interval Potential Survey (CIPS) and Direct Current Voltage Gradient (DCVG) is used along with in-line inspection to determine abnormal points along the right of way. The machine learning methods are explored in two aspects: 1) unsupervised clustering on regional environment and soil information for corrosivity similarity analysis; and 2) supervised classification for identifying corrosion defects and severities from the potential profile obtained from CIPS.

This subtask is to develop pipeline degradation prediction models based on the inspection and survey data, so that the evolution of defect quantities (e.g., corrosion, crack, mechanical damage, and interactive threats) can be predicted using the corresponding influencing factors. Bayesian neural network (BNN) will be deployed to learn the features from the datasets to propose defect generation and propagation models considering uncertainties. Therefore, the task will focus on using probabilistic models, BNN, and other machine learning methods to predict defect growth over time using in-line inspection data and influencing variables. Depending on the data obtained from the industry partner, the research team will explore different machine learning methods for prediction of degradation evolution depending on the dataset size, potential influential variables, linearity, training time, model accuracy, and output interpretability. After that, the predicted defects will be used to estimate the probability of pipe failure with respect to pipe capacity and support risk-based decision making.

* Task 5: Calculation of the expected cost of failure, how to determine the consequence of the major pipeline failures. Discussed the potential utilization of PHMSA annual reports, incident data, etc.

It was pointed out the PHMSA incident data could be potentially used to help quantify the pipeline failure impact. The PHMSA database includes: 1) PHMSA annual report, pipeline mileages and pipeline replacement updates; 2) PHMSA incident flagged files; 3) PHSMA pipeline incident data. The research team reviewed these PHMSA database and found that the incident reports by operators and the flagged files are believed to have the most relevance to the estimated incident cost. Four types of pipeline systems are included, natural gas transmission and gathering, natural gas distribution, hazardous liquid and carbon dioxide, and liquefied natural gas. The incident reports are compiled based on the pipeline system type and various time periods. The information collected in the incident reports include the location, incident area type, time, pipe information, operation information, failure type (e.g., leak, rupture etc.), shutdown time, injuries and fatalities, environmental impact (e.g., soil contamination, wild life impact), and estimate costs (property damage, emergency, etc.). These cost data will be analyzed for the potential use in the risk-based life-cycle cost analysis.

The expected cost of failure can be also collected from other resources and cost analysis tools in the literature. For example, Gomes et all. (2013) provided the pipeline failure costs due to small leak, burst, and rupture, which were derived based Canada pipeline incident reports. Other researchers (such as, Xie & Tian 2018; Sahraoui et al.; Abubakirov et al. 2020) also provided cost data based on typical industry estimation practices, considering human, environmental, and economic losses.

Abubakirov, R., Yang, M., & Khakzad, N. (2020). A risk-based approach to determination of optimal inspection intervals for buried oil pipelines. *Process Safety and Environmental Protection*, *134*, 95-107.

Gomes, W. J., Beck, A. T., & Haukaas, T. (2013). Optimal inspection planning for onshore pipelines subject to external corrosion. *Reliability Engineering & System Safety*, *118*, 18-27.

Sahraoui, Y., Khelif, R., & Chateauneuf, A. (2013). Maintenance planning under imperfect inspections of corroded pipelines. *International journal of pressure vessels and piping*, *104*, 76-82.

Xie, M., & Tian, Z. (2018). Risk-based pipeline re-assessment optimization considering corrosion defects. *Sustainable Cities and Society*, *38*, 746-757.

*Literature Review*

The research team conducted Task 1 Literature Review and summarized the review results related to Pipeline Defect Growth and Performance Modeling.

Firstly, the traditional prediction models that are widely used for pipeline degradation were introduced. Then, the artificial neural network (ANN) that have been applied to pipeline defects prediction in recent years was discussed. In addition, the Bayesian neural network (BNN) was included to explore its potential to be applied to pipeline prediction in the future. After that, strengths and limitations of these models were analyzed. Finally, some recommendations were given. Using machine learning methods to integrate datasets from different sources may solve the problems of insufficient data. Using Bayesian inference can consider the measurement uncertainties and improve the model performance. ANN model is useful for selecting the most critical factors and increasing the prediction accuracy in complex situations. BNN model may be a promising method for defect growth prediction because it has strong generalization ability and can automatically adjust network parameters.

In summary, the following conclusions and recommendations can be drawn from the detailed literature review.

1. Using machine learning methods such as transfer learning to integrate datasets from different sources and using Bayesian inference to consider the uncertainties of data may be possible methods to solve the problems of insufficient data and measurement uncertainties.
2. ANN models that can consider more factors may be useful for selecting the most critical factors and increasing the performance accuracy of prediction in complex situations.
3. AI-based models are effective for the pipeline degradation prediction. They can achieve satisfactory accuracy at certain cases. BNN model may be a promising AI-based method for defect growth prediction because it has strong generalization ability and can automatically adjust network parameters.

**Project Activities with Cost Share Partners:**

Cost share is provided by Rutgers University during this quarterly period as budgeted in the proposal.

**Project Activities with External Partners:**

The PI (Dr. Hao Wang) have conducted several virtual meetings with the Co-PI (Dr. Qindan Huang) at Marquette University to discuss the project work plan and literature review findings.

**Potential Project Risks:**

N/A

**Future Project Work:**

The research team will start Task 2 Data Collection from Existing Literature and Industry Partners. The purpose is to collect field inspection records from direct assessment and ILI and indirect survey/examination data from the industry partner and other available data sources.

**Potential Impacts to Pipeline Safety:**

The literature review summarized current practice and new developments in probabilistic performance degradation models of pipeline, which can aid pipeline operators better predict safety risk based on inspection data.