CAAP Annual Report

Date of Report: *10/09/2023*

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

Annual Period: **From** *(09, 30, 2022)* **to** *(09, 29, 2023)*

Contract Number: *693JK32250004CAAP*

Project Title: *Development of Compatibility Assessment Model for Existing Pipelines for Handling Hydrogen-Containing Natural Gas*

Prepared by: *Ramadan Ahmed*

Contact Info.: *r.ahmed@ou.edu, 405-473-7771*

Table of Contents

Section A: Business and Activities

(a) Contract Activities

- Contract Modifications:
    - Brady Dague was replaced by Mary McDaniel as the Agreement Officer Representative (AOR) for the cooperative agreement on November 14, 2022. Afterward, Mary McDaniel was replaced by Nusnin Akter on May 04, 2023. Currently, Nusnin Akter is the AOR for the cooperative agreement. Also, Jeffrey Luo was appointed as the Technical Task Initiator (TTI) for the agreement on May 04, 2023.
- Educational Activities:
    - Student mentoring: Two graduate and one undergraduate research assistants are currently being supervised by the Principal Investigators (PI and Co-PI). Students benefit from their guidance on study strategies, time management, energy transition, and hydrogen embrittlement challenges associated with repurposing natural gas pipelines.
    - Student internship: The PI and Co-PI assist students in exploring career options, setting goals, and developing skills and networks necessary for success in the energy field in the future. Career fairs and job search activities are encouraged for students to find internships during the summer. Additionally, PI and Co-PI assist students in developing self-confidence, resilience, and other life skills.
- Dissemination of Project Outcomes:
    - Journal Publications: Our research team diligently disseminates project outcomes and findings to a broad audience. One peer-reviewed research paper was prepared and submitted for publication in this reporting year. Two papers are also being prepared for publication in journals. The papers provided a detailed description of the research methodology, findings, and implications.
    - Conferences and Workshops: The project team prepared two posters for presentation at the PHMSA's R&D Forum (October 31-November 1, 2023). In the posters, findings and future plans are presented to industry experts and a wide range of stakeholders, such as members of academia, small businesses, and the general public.

(b) Financial Summary

- Federal Cost Activities: **Table 1** shows the expenses incurred in different budget categories as of September 30, 2023.
    - PI/Co-PIs/students involvement: Salaries and wages total $68,744.20, which includes graduate research assistant salaries and PI summer salaries. The project

has so far been conducted by four graduate research assistants who have performed a variety of tasks. A total of $27,936.60 is spent on tuition support for graduate students in this reporting year. Additionally, we have hired an undergraduate student to assist with experiments.

- Materials purchased/travel/contractual (consultants/subcontractors): The total expenditure related to supplies is $3624.28, which comprises expenses for test materials (vintage pipes), test gases (hydrogen, natural gas, and nitrogen), and miscellaneous supplies needed for running experiments. The total cost of fabricated equipment is $30,755.83, which includes items purchased to modify the experimental setup. These items include measuring instruments, hydraulic cylinders, a gas pressure booster, structural components for building the experimental setup, and heating jackets.

Table 1: Federal cost by budget category for Year 1

No.	Budget Categories	Cumulative Expenses
1	Salary & Wages	$ 68,744.20
2	Fringe Benefits	$ 12,092.40
3	Supplies	$ 3,624.28
4	Travel - Domestic	$ -
5	Other	$ 40.00
6	Equipment	$ 30,755.83
7	Tuition	$ 27,936.60
8	F&A 55%	$ 46,475.54
	Total	$ 189,668.85

- Cost Share Activities:

  - Cost share contribution: **Table 2** presents the University of Oklahoma (OU) cost share during the reporting period in each budget category. OU provided a total cost share of $63,654 in Year 1 in the form of personnel salary ($30,603), fringe benefits ($10,464), and overhead costs ($22,586). As part of the cost sharing, the PI and Co-PI participated in various research and development activities, including supervising research assistants and technical personnel, conducting hydrogen embrittlement research, and designing experimental setups.

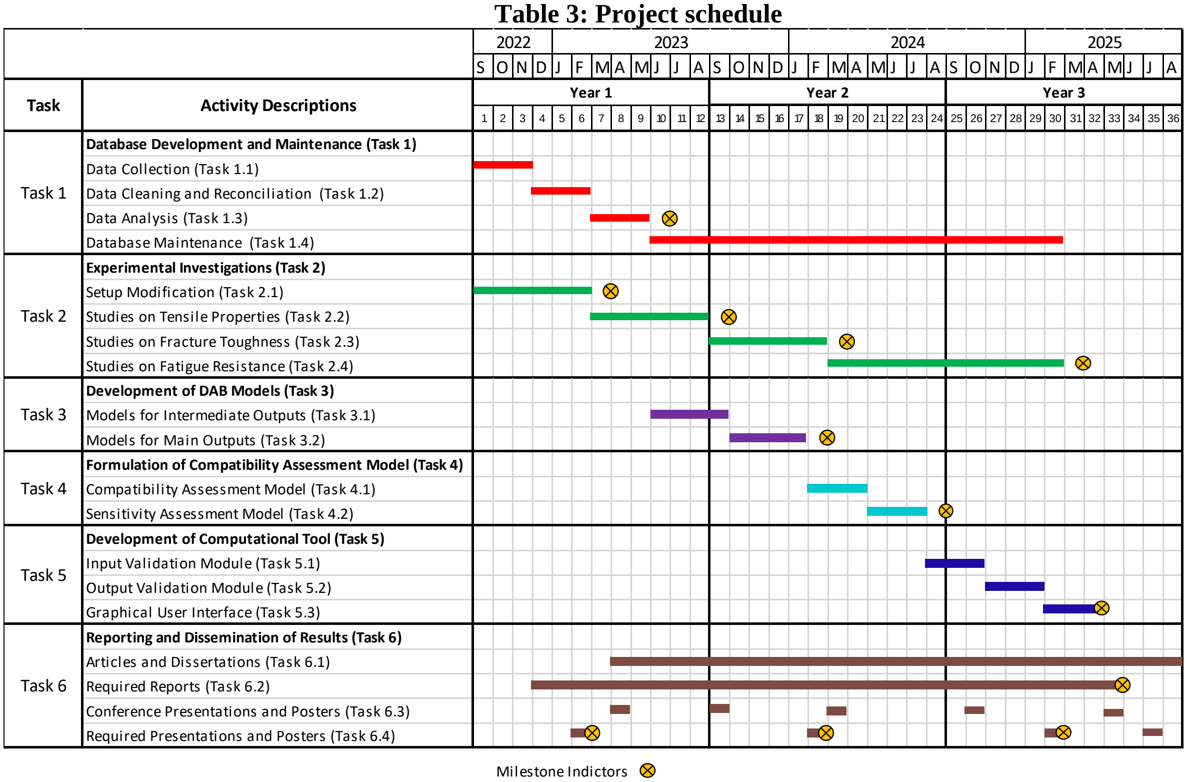
Table 2: OU cost share for Year 1

No.	Budget Category	OU Cost Share
1	Salaries and Wages	$ 30,603
2	Fringe Benefits	$ 10,464
3	Equipment	
4	Travel	
5	Materials and Supplies	
6	Tuition	
7	Indirect Costs	$ 22,586
	Total	**$ 63,654**

(c) Project Schedule Update

- Project Schedule: **Table 3** displays the sequence of project tasks and milestones for the project. We are progressing according to schedule with all theoretical and modeling tasks. However, the delay in manufacturing autoclaves required for building test setups has caused a 6-month delay in modifying experimental facilities (Task 2.1).

Table 3: Project schedule



- Corrective Actions: The manufacturer indicated a significant backlog of orders, so the manufacturing delay was anticipated earlier. As a result, an additional autoclave was ordered to perform parallel experiments once testing begins. Therefore, the delay will not significantly affect the overall project schedule since experimental studies (Tasks 2.2, 2.3 and 3.2) will be conducted in a shorter time frame. Additionally, spare parts are purchased for critical and custom-made items, such as hydrogen resistance strain gauges with extended lead time.

(d) Status Update of the 4th Quarter Technical Activities

- Task 1.4: Database Maintenance

Our ML models are being routinely maintained to improve their performance and accuracy. Periodic database maintenance began in Quarter 4. The database has been updated with a number of datasets recently published in articles. With the generation of original experimental data and the addition of data from other ongoing HE studies, our database will continue to grow. As part of the data maintenance process, the team cleaned up the data to identify and correct any inaccuracies or inconsistencies. Several datasets have been verified and corrected for inconsistencies based on their sources.

- Task 2.1: Setup Modification

The test facility had to be modified to measure hydrogen embrittlement in situ. The modification required purchasing autoclaves with accessories. A four-month delay was experienced in the manufacturing of the autoclaves. They were delivered at the end of Quarter 4. The autoclaves were then inserted into aluminum jackets to allow heating and cooling fluid to circulate. We tested each jacket for leaks and repaired any that were found. After that, they were mounted on test setups equipped with all the necessary instruments (**Fig. 1**). They were connected to the heating fluid lines. Before the placement of the autoclaved, load cells were calibrated using a crane scale. The calibration accuracy will be verified by applying a predetermined load on the cell with a hydraulic cylinder. The autoclave and other equipment handling test fluids are placed in a separate room where only intrinsically safe equipment and instruments are used. The room is equipped with hydrogen and natural gas detectors with alarm systems.

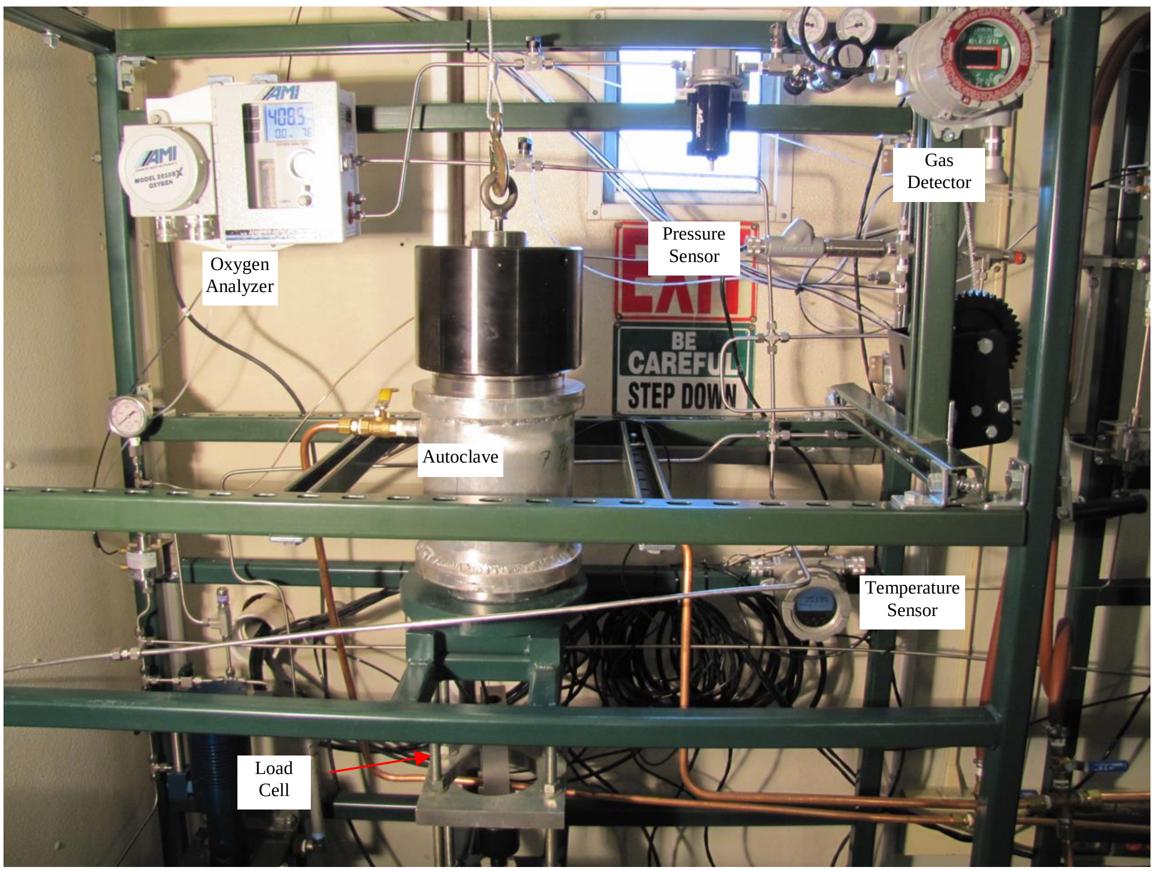


Fig. 1: Test setup with required components

Furthermore, the research team spent time calibrating the inline gas oxygen content analyzer (Fig. 1). The analyzer has a capability to measure up 10,000 ppm. But it is calibrated to measure low oxygen concentrations (0 to 100 ppm) in the gas stream. Hence it is very sensitive to detect the presence of small amount of oxygen in the autoclaves. The presence of oxygen in an autoclave, even in small amounts, can significantly influence measurements during the embrittlement test. Hence, the test requires meticulous air purging to remove oxygen from the environment. Due to this, each autoclave has an inlet for gas and an outlet for air purging to increase measurement reliability. Flexible tubes deliver gas to the bottom of the autoclave from the inlet line. As a result of the top outlet and bottom injection, stagnant air is efficiently displaced, and the autoclave is purged effectively. The autoclaves are connected to an oxygen analyzer on their outlet lines to monitor the oxygen content of the exit gas.

- Task 2.2: Studies on Tensile Properties

In year one, this task was supposed to be completed but was delayed because of Task 2.1. In the meantime, specimen and holder designs were developed to fabricate these items while waiting for setup modifications to be completed. The items have been manufactured and are ready to be used in the experiments. A complete set of specimens (a total of 450 pieces) for the project has been manufactured and received this quarter. Specimens (**Fig. 2**) are marked for their material type (X52, X60, and X70) and placed in vacuum-sealed bags. This task is rescheduled to be completed in the second quarter of year two.

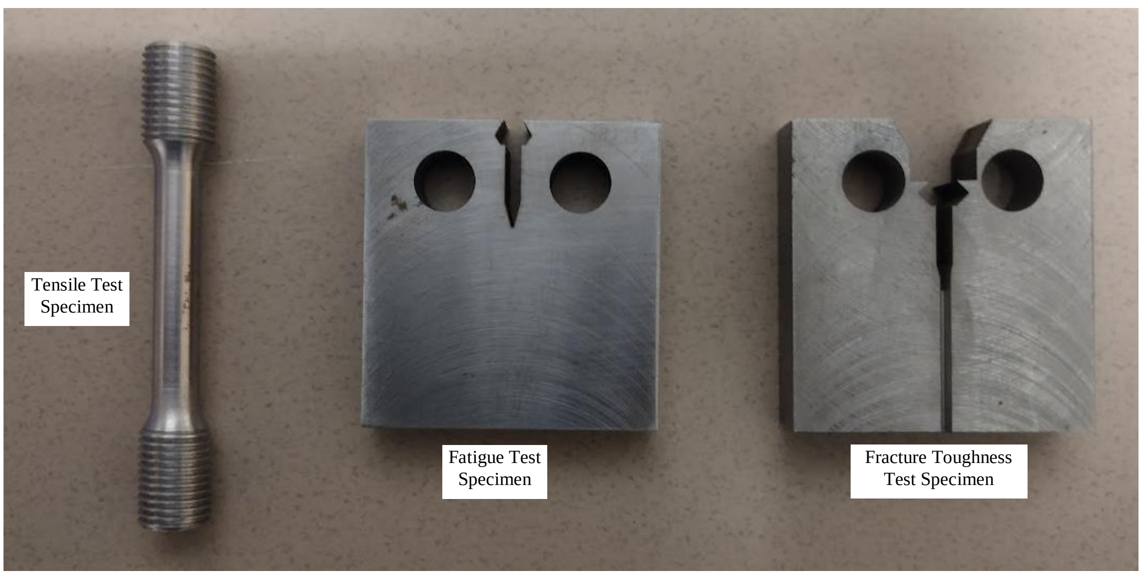


Fig. 2: Test specimens cut from vintage pipes

- Task 3.1: Models for Intermediate Outputs

Since the beginning of the fourth quarter, our research team made efforts to develop data analytics based (DAB) models for intermediate outputs such as reduction in area, hydrogen diffusivity, and solubility. By leveraging existing experimental data and employing supervised learning techniques, machine learning models have been developed to provide valuable insights into hydrogen behavior in carbon steels and contribute to developing the main output models that forecast pipeline materials suitable for hydrogen transportation. An article has be prepared and submitted showing the performance of reduction of area model. Similar models have been developed for maximum elongation of tensile specimen and hydrogen diffusivity and permeability.

Section B: Detailed Technical Results in the Report Period

1. Background and Objectives in the 1st Annual Report Period

Background

The purpose of this study is to investigate Hydrogen Embrittlement (HE) of pipeline materials and evaluate their suitability for handling blended and pure hydrogen gases. As a result of the complicated hydrogen permeation process and the various factors that affect embrittlement phenomena, existing HE models have limited application. In addition, the available data for evaluating the models in a variety of field conditions is inadequate. However, preliminary studies suggest that Data Analytics Based (DAB) models can be used to predict HE effects on mechanical properties. Consequently, HE failure can be predicted under various conditions by means of an integrated DAB model formulated correctly and tested extensively with different pipeline material types.

As part of this project, DAB modeling techniques are being developed to build a comprehensive compatibility assessment model for the use of existing pipelines to transport blended and pure hydrogen gas. A queryable database has been developed using publicly available experimental data to build the components of the DAB model. The database will also be expanded and enhanced by conducting experiments. Various data analytics methods have been evaluated for their suitability for predicting HE. The compatibility assessment model will include methods that demonstrate acceptable performance. A computational tool (software) based on the compatibility assessment model will be the main deliverable of the project. A pipeline's suitability to transport hydrogen or blended gases can be determined using the tool. In addition, it will determine whether HE-inhibiting molecules should be added to the gas processing operations or if adjustments must be made to the gas processing operations.

Objectives in the 1st Annual Report Period

The proposed 3-year (36-month) project which began on September 30th, 2022, and ends on September 29th, 2025. The objectives in the first year are: i) developing an HE database by collecting relevant data (Task 1.1), cleaning and preprocessing the collected data (Task 1.2), and performing data analysis through data observation, exploration, organization, and transformation (Task 1.3); ii modifying experimental setup by installing a new autoclave that allows in-situ hydrogen embrittlement experiment under high-pressure conditions (Task 2.1); iii) perform experimental studies to investigate the impact of hydrogen embrittlement on tensile properties of pipeline materials (Task 2.2); and iv) development of intermediate data analytic-based models that are needed for establishing main output models that are required for formulating the compatibility assessment model (Task 3.1).

2. Experimental Program in the 1st Annual Report Period

Experimental Design

An existing test setup with three test autoclaves has been modified to perform hydrogen embrittlement (HE) studies on different pipeline materials. Two newly constructed autoclaves (test cells) with a volume capacity of 3.1 L are installed in the setup for testing in-situ hydrogen embrittlement in pipeline steel (**Fig. 3**). The cells are large enough to accommodate various specimen designs (**Fig. 4**), holders, and clip-gauge assemblies. The cells are jacketed to allow heating and cooling by circulating heat transfer fluid (glycol) during the experiment. Custom-made aluminum jackets are designed and manufactured to cover the autoclave body as shown in **Fig. 5**. A high-capacity chiller and a 6-KW electric heater control the glycol temperature.

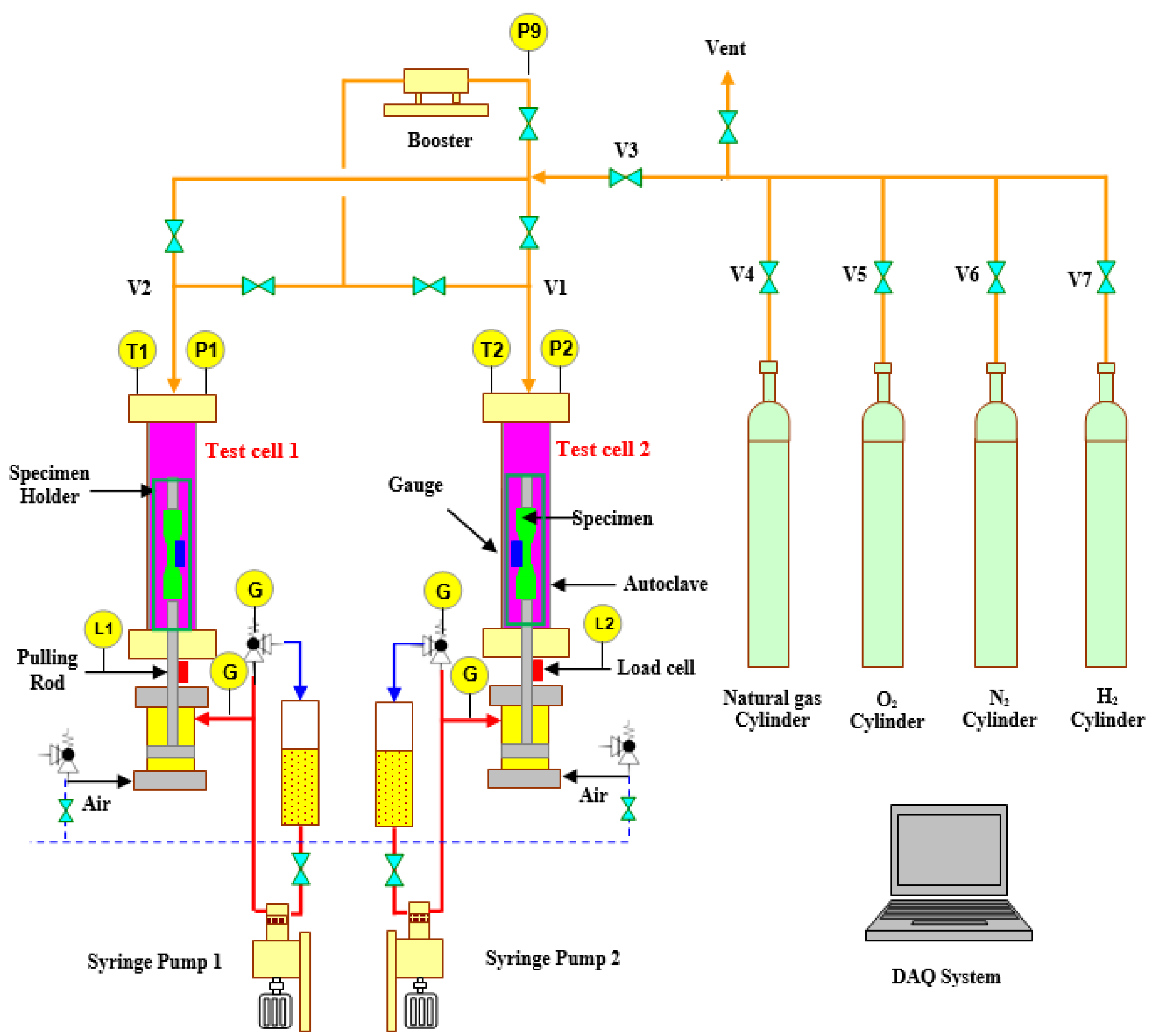
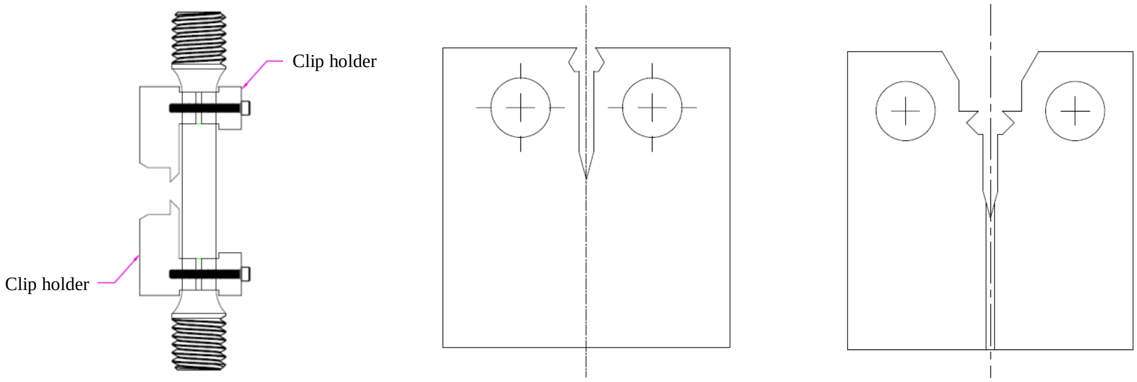


Fig. 3: Simplified schematic of test setup



(a) **(b)** **(c)**

Fig. 4: Specimen designs for different studies: a) tensile test; a) fatigue experiment; and c) fracture toughness test

(a) **(b)**

Fig. 5: Autoclave before (a) and after (b) covered with a heating jacket

The specimen holder allows the positioning of a test specimen in the cell and ensures its proper alignment during the test. Holders are made of aluminum tubes, with disks at the top and bottom. The top disk holds the upper side of the specimen during the experiment. The bottom disk has a hole in the center for passing a rod (pulling rod) attached to the specimen. The disk provides lateral support to the tube to maintain specimen orientation. Furthermore, the top disk has four thumbscrews to maintain stability when high loads are applied to the specimen. Also, four side holes (1.5 inches in diameter) are made on the wall of the tube to allow the effective purging of oxygen from the cell and easy assembly of a clip gauge.

Custom-built clip gauge (**Fig. 6**) measures strain rate during the test. The gauge is designed according to ASTM standard (E399 − 22). With its clipping design, it can be easily assembled with compact specimens (Figs. 4b and 4c) for fatigue and fracture toughness testing. For tension test specimens, the gauge requires an adaptor (clip holder) to attach them to specimens

(Fig. 4a). The clip holder has two jaws (upper and lower) that clamp to the specimen using screws. Hydrogen-resistant strain gauges are used in the gauge construction. A high-pressure wire-sealed gland designed for hazardous environments is installed on the autoclave lid to power the gauge and transmitter analog outputs. The gauge outputs are amplified to be monitored and recorded with the data acquisition system.

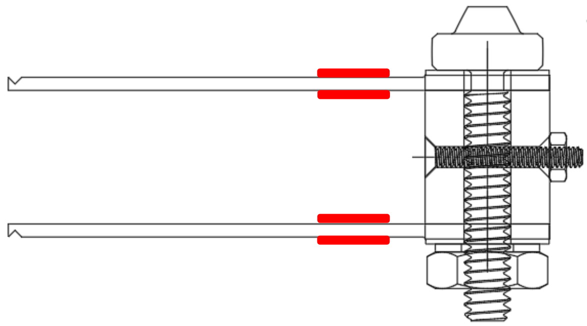


Fig. 6: Custom-built clip gauge

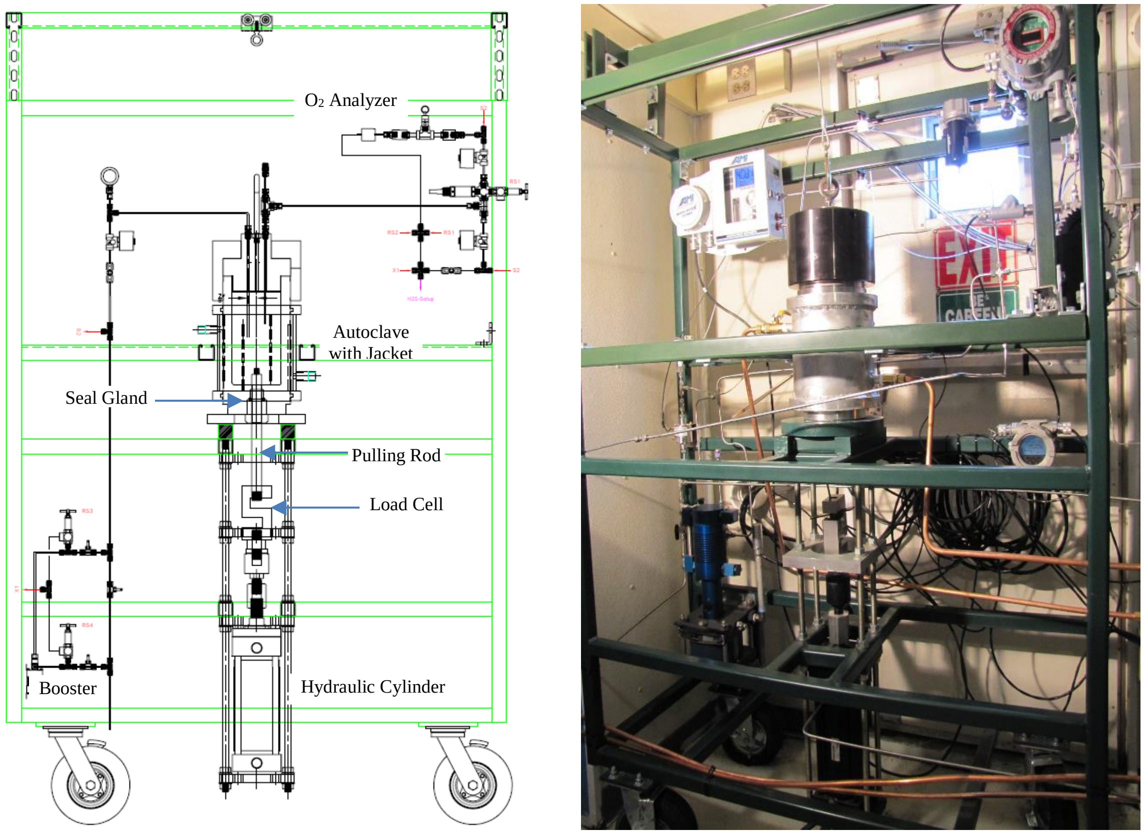
As the specimen is stretched during the test, an axial force is applied to the pulling rod using a hydraulic cylinder mounted underneath **(Fig. 7**). There is a seal gland on the bottom side of the autoclave for the rod. It connects directly the specimen to a load cell, which is connected to the hydraulic cylinder through a universal joint. When pressured at 3000 psi, the cylinder generates a maximum force of 25,000 pounds. The load cell can measure up to 15,000 lbs. A programmable syringe pump controls hydraulic pressure during the experiment. In addition to analog input, the pump allows the user to manipulate the output pressure during fatigue tests.

Air purging is a critical aspect of the embrittlement test, as the gas oxygen content significantly influences measurements. Hence, to improve measurements reliability, each autoclave has one gas inlet and one outlet to allow efficient air purging. The inlet line is equipped with flexible tubes that deliver gas to the autoclave bottom. With the top outlet, the bottom injection ensures the effective displacement of stagnant air and efficient purging of the autoclave. An oxygen sensor is installed on the outlet line of the autoclaves to monitor exit-gas oxygen content and confirm proper purging of the air from the system.

In the gas supply system, various gases such as oxygen, nitrogen, natural gas, and hydrogen are stored and supplied during the test. The vendor supplies the gases in high volume capacity (261 and 345 SCF) cylinders at pressures of 2000 to 2500 psi. On the inlet line of the cylinder, a gas booster is installed to increase the gas pressure. This is when the supply cylinder pressure is less than the desired test pressure.

Various sensors are installed in the setup to measure and monitor critical test parameters such as temperature and pressure. To measure temperature, a thermocouple is inserted into the reactor through a compression fitting. Pressure at the autoclave outlet is measured with a pressure transducer. In addition, lower explosive limit (LEL) sensors are mounted on the setup to detect hydrogen and natural gas leaks. All components of the setup in contact with the test fluid are made of stainless steel 316. The autoclave and other equipment handling test fluid are placed in a separate room where only intrinsically safe equipment and instruments are used.

The room is equipped with hydrogen and natural gas detectors with alarm systems.



(a) (b)

Fig. 7: Design (a) and photo (b) of experimental setup

Test Procedure

To perform the experimental investigation, commonly used grades (X52, X60, and X70) of vintage pipes were obtained from the Pipeline Research Council International. Pipes were thick enough to cut flat and cylindrical specimens. One hundred fifty specimens have been manufactured for the three specimen types shown in Fig. 4, 50 pieces for each material type. The specimens are stored in a dry vacuum environment to prevent surface alteration due to surface corrosion (rusting). All specimens are marked to identify their material type unmistakably and prevent human error in measurement reporting.

Testing of the specimens involves several steps, including pre-cracking, cleaning, and preparation, placing the specimen in the autoclave, purging of the autoclave to remove air, adjusting the temperature of the autoclave, pressurizing the cell, charging hydrogen, applying the desired load and measuring the strain as a function of time until the experiment ends, removing the specimen from the autoclave, and examining it microscopically.

1. **Pre-cracking:** Fatigue pre-cracking is needed for CT specimens used for fracture toughness and fatigue strength experiments. The procedure is performed under atmospheric conditions by applying a fluctuating load at a constant load ratio and monitoring the crack opening using conductivity measurements.

2. **Cleaning and Preparation:** The specimens must be clean and have an equal smoothness level on the surface so that the measurements can be consistent. Standard surface finishes are applied to specimens manufactured from pipe samples. A vacuum seal is applied to the containers in which they are stored. Testing samples is performed after they have been degreased and cleaned with acetone.

3. **Placement of Specimen in Autoclave:** As a first step in mounting the specimen in the autoclave, the lid (cover) of the autoclave is removed, and the specimen holder is raised using a hydraulic cylinder attached to the pulling rod. The cylinder is pneumatically actuated. Afterward, the specimen is placed on the holder, and the clip gauge is attached to the specimen. Then, the holder components (top disk and thumbscrews) are assembled. The assembly is then lowered by retracting the hydraulic cylinder. Gas injection and distribution tubes are connected to the autoclave lid. The autoclave is then sealed by placing the cover and screwing its cap. Gas inlet and outlet lines are then connected. A thermocouple is inserted into the cell.

4. **Autoclave Purging and Temperature Adjustment:** Gas purging and autoclave temperature adjustment are done simultaneously. Heating or cooling medium is circulated through the autoclave jacket to adjust the temperature in the autoclave. Purging is performed step by step using test gases (natural gas or hydrogen). Temperature and oxygen content in the autoclave are monitored during purging. The process is stopped once the desired temperature and oxygen content in the autoclave are established.

5. **Cell Pressurization:** Cell pressurization is performed stepwise to control gas composition accurately. Hence, the test gases are injected into the autoclave according to their molar fraction. This means the gas with the lowest molar fraction is injected first, and the gas with the highest is injected last.

6. **Hydrogen Charging:** The specimen in the autoclave is aged for 2 hours after pressurization. Temperature and pressure are monitored and maintained at the desired level. A data acquisition computer automatically controls temperature, and pressure is manually adjusted.

7. **Testing:** After charging, the desired load is applied, and the resulting strain is measured as a function of time until the end of the experiment.

8. **Specimen Recovering:** Upon completion of the test, the autoclave is opened to collect and examine the specimen.

9. **Microscopic Evaluation:** The specimen is examined using an optical microscope to evaluate crack characteristics, failure type, assess necking behavior, and determine area reduction.

3. Results and Discussions

Task 1: Database Development and Maintenance

This task consists of four subtasks (Table 3): data collection (Task 1.1), data cleaning and reconciliation (Task 1.2), data analysis (Task 1.3), and database maintenance (Task 1.4). In the reporting year, three of these tasks were completed. Database maintenance will continue into the third year of the project.

- Task 1.1: Data Collection

The data collection was completed during this reporting year's first quarter. To accomplish this task objective, an extensive literature survey was conducted to establish a comprehensive hydrogen embrittlement (HE) database of common pipeline materials used in natural gas transportation and other industrial operations. The survey focused on HE studies involving direct pressure charging of materials in a hydrogen environment. Data from various experimental studies, including tensile testing, fracture toughness experiments, and fatigue testing, were gathered to create a master database. Measurements that are often used for assessing HE, such as area reduction, maximum elongation, fracture toughness, and fatigue crack growth rate, are included in the database.

- Task 1.2: Data Cleaning and Reconciliation

Data cleaning and reconciliation was conducted to identify and resolve inaccuracies and inconsistencies in the master database. Most data collected from the open literature were in various formats; some were not in their original form, and others missed relevant information such as material composition and mechanical properties. The actual data sources were identified and obtained to recover original data sets reported in secondary references as regression lines. The database was then cleaned by replacing digitized datasets from regression lines with actual datasets. Following that, we resolved the missing data issues by imputation or removing the data sets. Material composition and mechanical properties presented in various standards were used to input missing information. The imputation was mostly used in cases with limited incomplete data, such as lack of material composition or mechanical properties (yield strength, hardness, and ultimate strength).

Furthermore, we used available machine learning techniques to resolve missing data issues by imputation or removing the data sets. For instance, **Table 4** shows summary statistics for the transformed variables in the fatigue dataset prepared for data analysis. During data processing, a data quality report was created. The dataset showed incomplete values of Ultimate Tensile Strength (Sx). The predictive mean matching (PMM) method was applied to forecast these missing values. Moreover, missing values for the Heat Treatment parameter were handled by creating another factor level called "Unknown" since they represent data points not reported in the original data sources.

Table 4: Summary statistics for variables in fatigue dataset

Variable	N	Mean	Std. Dev.	Min	25%	75%	Max
Metal Fe content	1360	98.466	0.291	97.8	98.3	98.706	99.312
Metal C content	1360	0.214	0.19	0.048	0.12	0.26	0.85
Metal Mn content	1360	1.121	0.284	0.47	0.82	1.29	1.53
Metal P content	1360	0.012	0.005	0	0.008	0.014	0.02
Metal S content	1360	0.019	0.011	0	0.012	0.026	0.042
Metal Si content	1360	0.163	0.098	0	0.11	0.25	0.31
Gas hydrogen pressure	1360	0.695	0.459	0	0	1	1
Gas CO2 content	1360	0.001	0.004	0	0	0	0.02
Gas SO2 content	1360	0.001	0.004	0	0	0	0.02
Gas O2 content (ppm)	1360	3.382	18.084	0	0	0	100
Frequency	1360	1.146	1.198	0.1	1	1	5
Load ratio	1360	0.258	0.235	0.05	0.1	0.5	0.8
ΔK	1360	2.936	0.453	1.716	2.635	3.255	4.538
Sy	1360	6.056	0.203	5.595	5.903	6.159	6.461
Sx	1360	610.29	87.287	463	526	675	835
da/dN	1360	-9.301	2.558	-20.592	-10.668	-7.566	-4.063

- Task 1.3: Data Analysis

This analysis was performed through observation, exploration, organization, and transformation to detect patterns, or regularities in data. Utilizing data analytics methods, we explored hidden patterns and relationships. For instance, the interdependency amongst various influential factors were analyzed and visualized using the cross-correlation matrix plots. **Figure 8** shows the effects of various materials properties and environmental parameters on the reduction area measured during tensile testing. The plot reveals a strong positive correlation between ultimate strength and yield strength, indicating that one parameter can essentially echo the other. Hence, to minimize the dimension of input-features, the yield strength was omitted. Conversely, alloying elements Cr, Ni, Mo showcased a strong negative correlation with Fe, leading to their exclusion from the input parameters. The presence of highly correlated parameters can lead to multicollinearity, a situation where two or more independent variables in a dataset share high correlation. This can result in overfitting and difficulty in discerning the unique impact of each correlated feature on the target variable. Thus, it hinders the determination of the feature that contributes most to the prediction. Following these adjustments, the final selection of input features for the development of the ML-model was whittled down to ten parameters, namely (Fe, C, Mn, P, S, Si, Al, Heat treatment, Pressure, Ultimate strength).

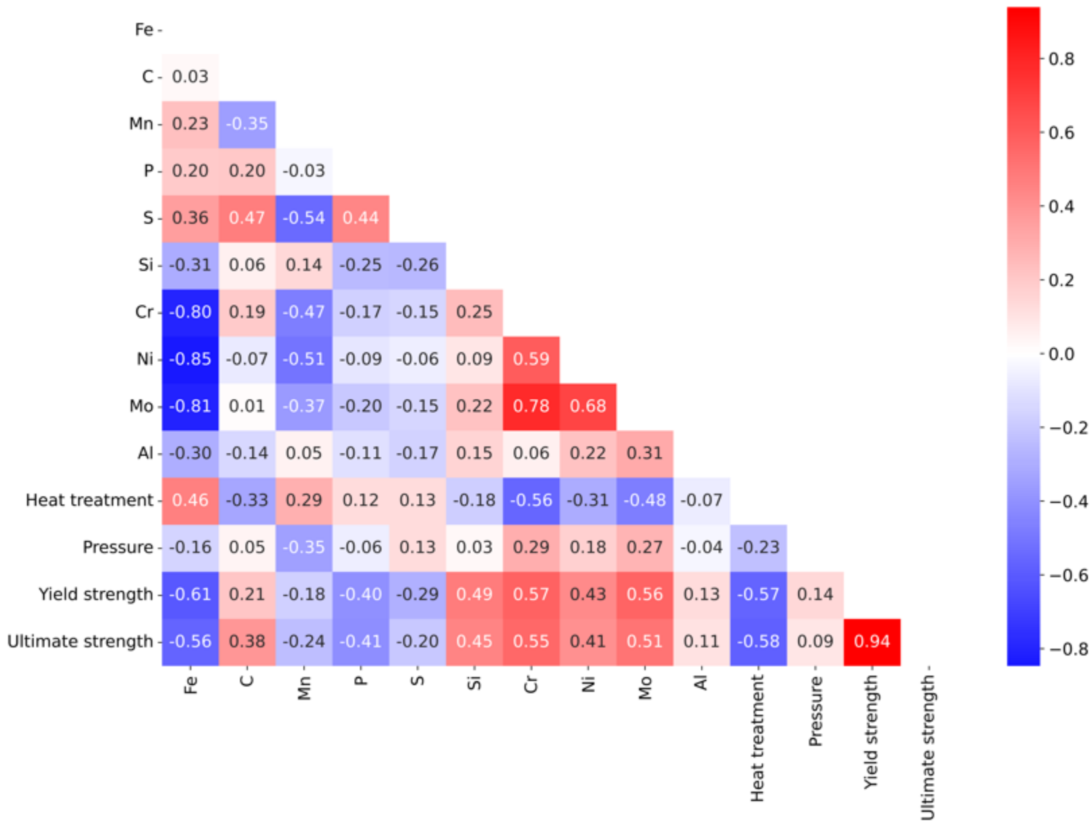


Fig. 8: Cross-correlation matrix plot of the selected parameters

- <u>Task 1.4: Database Maintenance</u>

In order to improve the performance and accuracy of our ML models, we have begun periodic database maintenance. As we generate new experimental data in this project and add data from other new HE studies, we will expand the database. We will also perform periodic data cleaning to identify and correct inaccuracies and inconsistencies. The revision history of the database is properly labeled after periodic maintenance. In case of data loss, corruption, or deletion, backup copies will be created to restore original data.

Task 2: Experimental Investigations

Four experimental subtasks consist of setup modification (Task 2.1), studies on tensile properties (Task 2.2), studies on fracture toughness (Task 2.3), and studies on fatigue resistance (Task 2.4) need to be completed as part of this task (Table 3).

- <u>Task 2.1: Setup Modification</u>

To perform hydrogen embrittlement measurements under in-situ conditions, the existing experimental setup (**Fig. 9**) had to be modified. Due to a delayed manufacturing process for autoclaves used to test specimens under pressure, the task was not completed on time. They were delivered in late September 2023. Once received, they were inserted into aluminum jackets that allow heating and cooling fluid to circulate. Each jacket was tested for leaks, and all leaks were repaired. After that, they were mounted on test setups that are equipped with all the necessary instruments. The autoclave and other equipment handling test fluids are placed in a separate room where only intrinsically safe equipment and instruments are used. The room is equipped with hydrogen and natural gas detectors with alarm systems. Currently, the autoclave system is being tested for planned experiments.

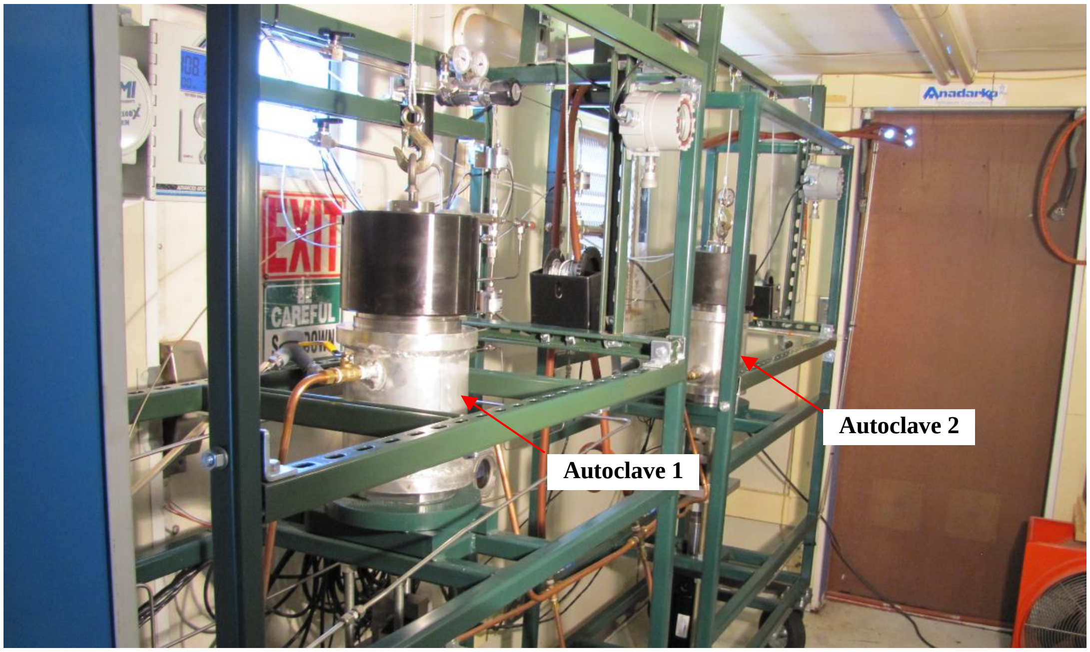


Fig. 9: Experimental setups for testing hydrogen embrittlement

- <u>Task 2.2: Studies on Tensile Properties</u>

This task aims to investigate the impact of HE on the tensile properties (yield strengths, ultimate tensile strength, reduction of area, and plastic strain before failure) of steel employed in pipelines. For these experiments, we use smooth specimens cut from thick vintage pipes. This task was expected to be completed in year one but was delayed because of the Task 2.1 delay. However, specimen and holder designs (**Fig. 10**) were developed to manufacture these items while waiting for setup modifications. The items have been manufactured and are ready to be used in the experiments.

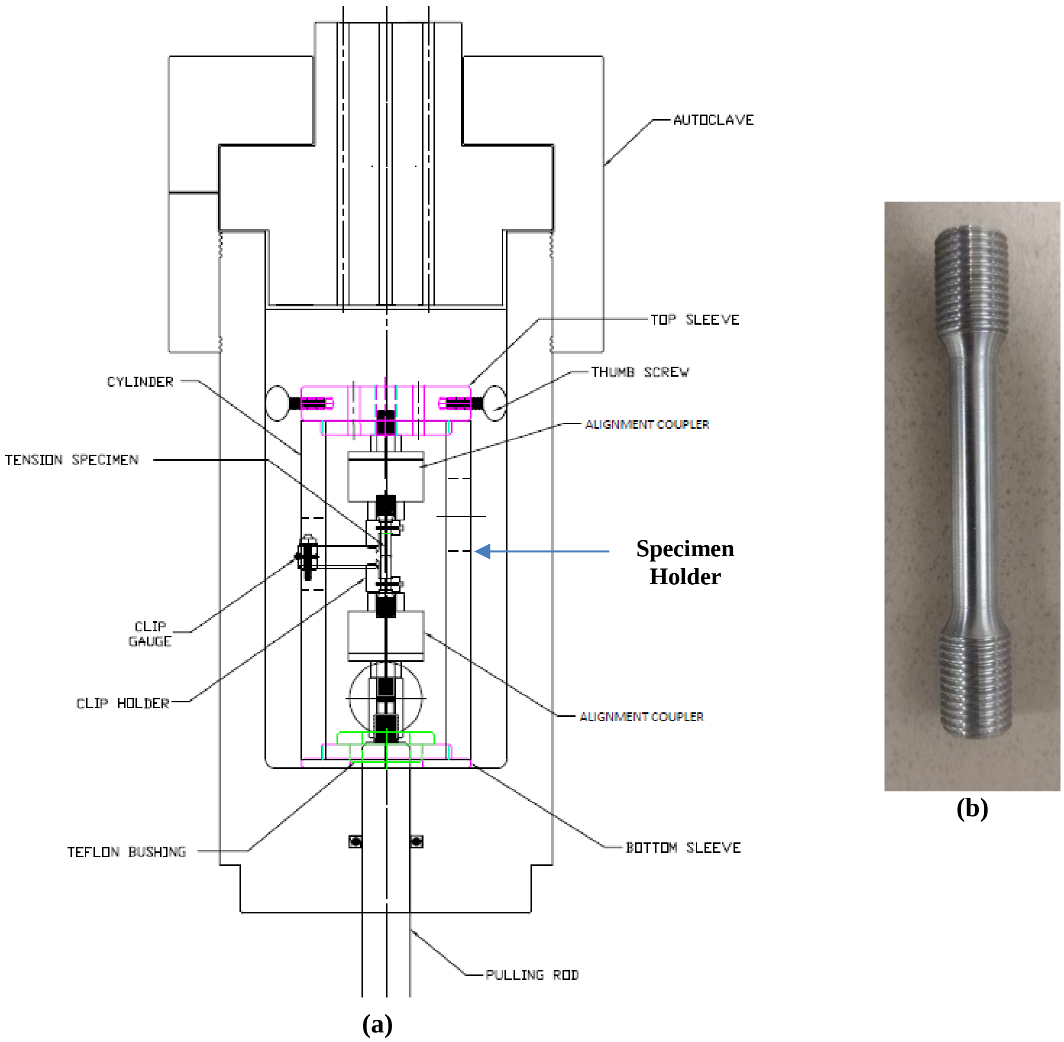


Fig. 10. Assembly design of tensile test specimen with holder in autoclave (a) and photo of tensile specimen (b)

Task 3: Development of DAB Models

This task aims to establish data analytics-based (DAB) models for main and intermediate outputs. It consists of model development for the intermediate (Task 3.1) and main (Task 3.2) outputs.

- <u>Task 3.1: Models for Intermediate Outputs</u>

Since the beginning of the fourth quarter, efforts have been made to develop DAB models for intermediate outputs such as reduction in area, hydrogen diffusivity, and solubility. By leveraging existing experimental data and employing supervised learning techniques, machine learning models have been developed to provide valuable insights into hydrogen behavior in carbon steels and contribute to developing the main output models that forecast pipeline materials suitable for hydrogen transportation.

Modeling Reduction of Area

To develop a reduction of area (necking) model, the dataset has been meticulously extracted from authoritative sources, including peer-reviewed journals, conference proceedings, and research papers from esteemed organizations [1-15]. The number of datasets gathered from each source is presented in **Table 5**. The focus is on low carbon and low alloy steel materials, with the gathered data comprising results of tensile tests performed under various conditions. The data collection process only considers low carbon and low alloy steels with smooth surfaces, excluding those with notched surfaces due to significant gaps in data and their limited abundance. The experimental conditions under which the tensile tests were conducted encompass environments of air, helium, and high-pressure hydrogen gas. The database is structured around multiple features, including mechanical properties, the chemical makeup of the materials, and the environmental conditions of the tests. The database is made up of tensile test data from 47 different low-carbon and low-alloy steels. The clean, ready-to-use dataset includes results from 236 tensile tests and encapsulates 22 distinct testing features.

Table 5: Data sources and corresponding number of collected data points

Data Source	Number of Data Points	Test Conditions	
		Strain rate (s^{-1})	Hydrogen Partial Pressure (MPa)
San Marchi and Somerday [8]	05	0.0001	6.9
Jewett, Walter [18]	10	0.000033	0.1-69
Wachob and Nelson [19]	03	0.0003	0.1
Hoover, Robinson [20]	13	0.0003 – 0.008	0.1 – 34.5
Loginow and Phelps [21]	22	0.00003-0.0003	0.1-97
Walter and Chandler [22]	25	0.000033- 0.0003	0.1-69
Walter and Chandler [23]	01	0.000033	69
Cialone and Holbrook [24]	21	0.0000021-0.001	0.1-6.9
An, Zhang [25]	04	0.0002	0.1-0.6
Clark and Landes [26]	01	0.0003	6.9
Padmanabhan and Wood [27]	01	0.0003	0.1
Bandyopadhyay, Kameda [28]	05	0.00003-0.0003	0.1
Robinson [29]	03	0.0003	0.1-4.14
Hoover [30]	54	0.0003	0.1-6.9
Takeda and McMahon [31]	01	0.0003	0.1
Walter and Chandler [32]	02	0.0003	0.1
Campbell [33]	09	0.000033	69
Hoover, Iannucci [34]	18	0.0003	0.1-6.9
Keller, Somerday [35]	06	0.0001	0.1-14
Ningileri, Boggess [36]	31	0.0001-0.00001	5.50-20.68
Fukuyama and Yokogawa [37]	01	0.0003	6.9

Seven DAB model techniques are considered for a comparative analysis in this task. These techniques included Decision Tree (DT), Random Forest (RF), Adaptive Boosting (AdaBoost), Gradient Boosting (GB), Extreme Gradient Boosting (XGBoost), Categorical

Boosting (CatBoost), and artificial Neuron Network (ANN). For conciseness, only the best-performing ML model - the CatBoost model - is described in detail here. The best model is determined by calculating the average of several results: the coefficients of determination for training and testing, MSE, RMSE, MAE, and the relative ranking of feature importance (as shown in **Table 6**). The model that produced results closest to this average was defined as the best-performing model.

Table 6: Performance evaluation and optimum model selection based on several metrics

Model Name	Coefficient of Determination		MSE	RMSE	MAE	Pressure	Su	Fe	C	Si	S	P	Al	Mn	HT*
	Training R^2	Test R^2													
RF	0.77	0.70	90.18	9.50	7.26	1	2	4	5	3	6	7	10	9	8
DT	0.73	0.69	95.31	9.76	7.45	1	2	4	7	8	9	3	10	5	6
XGBoost	0.78	0.74	78.05	8.83	7.07	1	5	6	9	2	3	4	10	7	8
ANN	0.75	0.65	108.04	10.39	8.18	1	2	4	7	10	9	3	5	8	6
AdaBoost	0.75	0.71	88.14	9.38	7.04	1	2	4	6	3	5	7	10	9	8
GB	0.74	0.71	88.03	9.38	7.29	1	2	3	6	4	7	5	10	9	8
CatBoost	**0.78**	**0.73**	**83.78**	**9.15**	**7.32**	**1**	**2**	**3**	**6**	**5**	**8**	**4**	**10**	**9**	**7**
Average	0.76	0.70	90.07	9.48	7.38	1.0	2.4	4.0	6.6	5.0	6.7	4.7	9.3	8.0	7.3

*Heat treatment

CatBoost is a member of the Gradient Boosted Decision Tree ML algorithm, which can handle both categorical and heterogeneous data. The method makes use of a specific type of decision tree known as an oblivious tree [16]. Unlike a regular decision tree, every node at the same level in an oblivious tree makes the same decision. Such trees are balanced, less prone to overfitting, and allow significantly faster execution at testing time [17]. In this task, the CatBoost model is trained with 80% of the dataset and then tested this model capacity to predict the ductility of the low carbon and low alloy steel with 20% of the data. For this algorithm, we do not need to perform data preprocessing; it can manage categorical data efficiently and predict without overfitting. CatBoost is robust to outliers and has a different scale of features. This supervised regression model is aimed to predict the HE susceptibility of particular low carbon and low alloy steel in terms of ductility reduction (reduction of area prediction) using the hydrogen pressure, materials' chemical and mechanical properties, and thermal conditions.

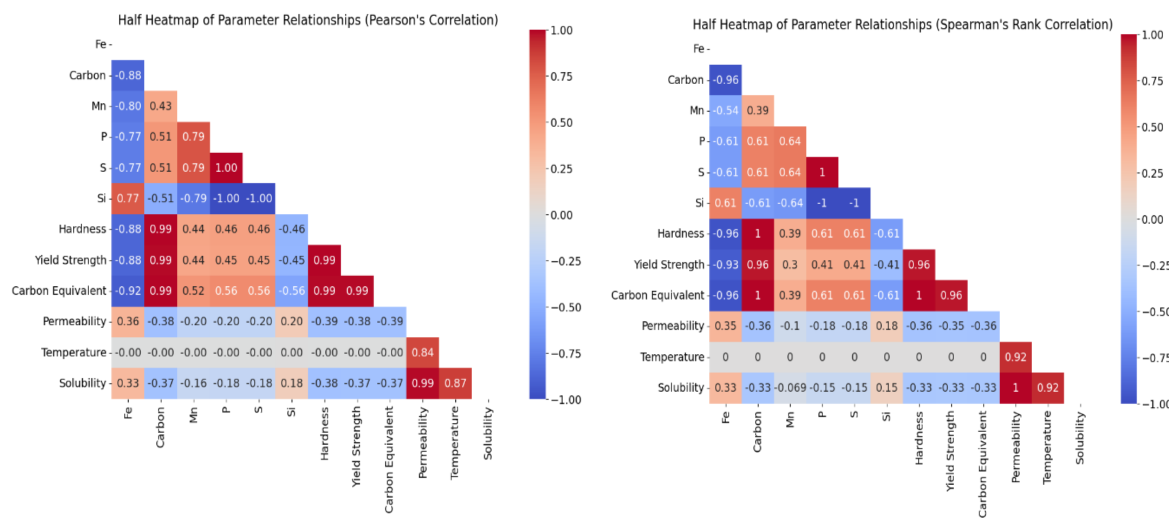
Modeling Hydrogen Diffusivity and Solubility

The datasets utilized in the diffusivity and solubility models have been compiled by conducting an extensive review of the existing literature. The majority of the data is generated from graphical representations found in these research papers, which are then converted into a usable format using digitization software. These datasets encompass seven distinct grades of carbon steels, along with their respective compositions, hardness levels, carbon equivalence, test environments, and output parameters (permeability, diffusivity, solubility). The database is being updated to ensure a comprehensive analysis by adding newly collected data for hydrogen solubility and diffusivity in carbon steels. A general overview of the dataset is presented in **Table 7**.

Table 7: Overview of diffusivity, solubility dataset

Carbon steel grades	Compositions	Temperature	Pressure	Total Datasets	References
Armco Ingot Iron, 1010, 1020, 1035, 1050, 1065, 1095	Fe, C, Si, S, P, Mn	500-900 K	0.01-0.7 MPa	2800	(Ghayedi and Khosravi, 2020)

Figure 11 shows a heat-map correlation between influential factors and hydrogen solubility. Pearson and Spearman's Rank correlation methods identify significant relationships among the input parameters. Specifically, iron (Fe) and carbon show a strong correlation (above 95%) with the carbon equivalent. Additionally, hardness is closely associated with both carbon equivalent and yield strength. Similarly, permeability is closely tied to solubility. Notably, phosphorus (P) also shows a high correlation with silicon (Si) and sulfur (S). In machine learning models, including highly correlated variables can lead to the issue of multicollinearity. Given that similar high correlations are observed with Spearman's rank correlation, it is decided to exclude carbon equivalent, hardness, permeability, Silicon, and Sulfur from the input variables for better model performance.



(a) **(b)**

Fig. 11: Pearson and Spearman's Rank correlation between features of hydrogen solubility prediction

Analysis of the importance of the feature from a random forest regression model for predicting hydrogen gas solubility in carbon steel reveals that temperature is the predominant factor, accounting for 79.8% significance (**Fig. 12**). This observation underscores a strong relationship between temperature and hydrogen solubility, emphasizing temperature's crucial role in mediating the physicochemical interactions between hydrogen and the metal. This observation is based on data collected at high temperatures (500 to 900°K). Low temperature solubility data are scarce, and we are looking for new studies. Composition variations also have a notable impact, suggesting that alterations in metal content can influence solubility. Iron and carbon, fundamental components of carbon steel, contribute with the importance of 10.1% and 5.2%, respectively, highlighting their roles in shaping steel's microstructure and properties. Yield strength, representing mechanical properties, registers importance of 4.9%, suggesting a potential correlation between the steel's mechanical characteristics and its hydrogen interactions. Conversely, features like Mn and P have no bearing on the model's predictions, each registering a 0% importance.

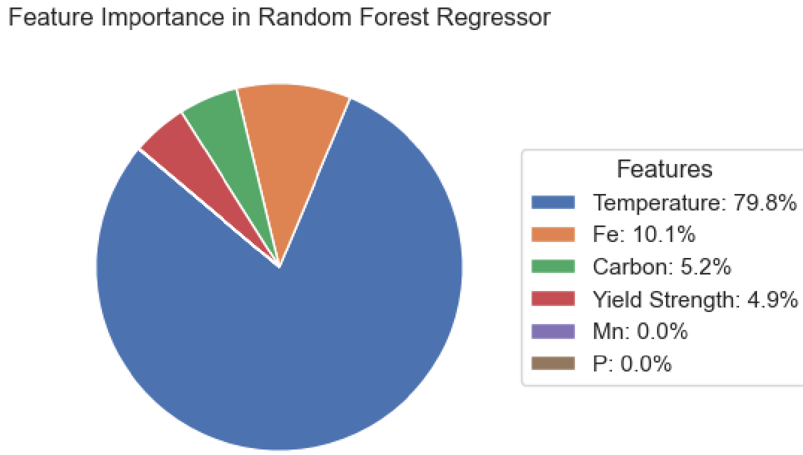


Fig. 12: Feature importance for hydrogen solubility prediction in random forest regression

4. Future work

Year two will include various activities, including conducting experiments, formulating different models, and developing a computational tool. Further, the project team will focus on completing delayed tasks (conducting tensile test experiments) to keep the project on track and complete the remaining tasks by its end date. Consequently, we plan to undertake the following research and development activities in the coming year:

1. Continuing database enrichment and maintenance (Task 1.4),
2. Performing experimental studies on tensile (Task 2.2) and fracture toughness (Task 2.3) of pipeline materials in parallel,
3. Starting experimental studies on the fatigue resistance of pipeline materials (Task 2.4),
4. Developing intermediate and main DAB models (Task 3),
5. Formulating a compatibility assessment model (Task 4), and
6. Starting the development of a computational tool (Task 5).

References

1. San Marchi CW, Somerday BP. Technical reference for hydrogen compatibility of materials. Sandia National Laboratories Albuquerque, NM, and Livermore, CA; 2012.
2. Jewett RP, Walter RJ, Chandler WT, Frohmberg RP. Hydrogen environment embrittlement of metals. United States: Rocketdyne International Corp., Canoga Park, CA; 1973.
3. Wachob HF, Nelson HG. Influence of microstructure on the fatigue crack growth of A516 in hydrogen. In: Bernstein IM, Thompson AW, editors. Proceedings of the third international conference on effect of hydrogen on behavior of materials. Warrendale, PA: The Metallurgical Society of AIME; 1980. p. 703-11.
4. Hoover W, Robinson S, Stoltz R, Spingarn J. Hydrogen compatibility of structural materials for energy storage and transmission. Final report. Sandia National Labs., Livermore, CA (USA); 1981.
5. Loginow AW, Phelps EH. Steels for Seamless Hydrogen Pressure Vessels. Journal of Engineering for Industry. 1975;97:274-82. https://doi.org/10.1115/1.3438550
6. Walter RJ, Chandler WT. Effects of high pressure hydrogen on metals at ambient temperature. United States: Rocketdyne International Corp., Canoga Park, CA; 1969.
7. Walter R, Chandler W. Influence of gaseous hydrogen on metals final report. Marshall Space Flight Center AL: NASA; 1973.
8. Cialone HJ, Holbrook JH. Sensitivity of Steels to Degradation in Gaseous Hydrogen. In: Raymond L, editor.^editors. Hydrogen Embrittlement: Prevention and Control. West Conshohocken, PA: ASTM International; 1988. p. 134-52. https://doi.org/10.1520/STP45297S
9. An T, Zhang S, Feng M, Luo B, Zheng S, Chen L, et al. Synergistic action of hydrogen gas and weld defects on fracture toughness of X80 pipeline steel. International Journal of Fatigue. 2019;120:23-32. https://doi.org/10.1016/j.ijfatigue.2018.10.021
10. Clark WG, Landes JD. An Evaluation of Rising Load KIscc Testing. In: Craig HL, editor.^editors. Stress Corrosion—New Approaches. West Conshohocken, PA: ASTM International; 1976. p. 108-27. https://doi.org/10.1520/STP28674S
11. Padmanabhan R, Wood WE. Hydrogen induced cracking in a low alloy steel. Metallurgical Transactions A. 1983;14:2347-56. https://doi.org/10.1007/BF02663310
12. Bandyopadhyay N, Kameda J, McMahon CJ. Hydrogen-induced cracking in 4340-type steel: Effects of composition, yield strength, and H2 pressure. Metallurgical Transactions A. 1983;14:881-8. https://doi.org/10.1007/BF02644292
13. Robinson SL. Hydrogen compatibility of structural materials for energy storage and transmission applications. United States: Sandia National Lab., Albuquerque, NM; 1976.
14. Hoover WR. Hydrogen compatibility of structural materials for energy storage and transmission. Annual report, October 1, 1977--September 30, 1978. United States: Sandia National Lab., Livermore, CA; 1978.

15. Takeda Y, McMahon CJ. Strain controlled vs stress controlled hydrogen induced fracture in a quenched and tempered steel. Metallurgical Transactions A. 1981;12:1255-66. https://doi.org/10.1007/BF02642339
16. Hancock JT, Khoshgoftaar TM. CatBoost for big data: an interdisciplinary review. Journal of Big Data. 2020;7:94. https://doi.org/10.1186/s40537-020-00369-8
17. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. Advances in neural information processing systems. 2018;31.