CAAP Quarterly Report

04/04/2025

Project Name: "Accelerating Transition towards Sustainable, Precise, Reliable Hydrogen Infrastructure (Super-H2): Holistic Risk Assessment, Mitigation Measures, and Decision Support Platforms"

Contract Number: 693JK32250007CAAP

Prime University: North Dakota State University

Prepared By: Xinyu Wang (PI), Zhibin Lin, Kevin Wang, Jace Anderson, Hong Pan, Mohsin Ali Khan.

Contact Information: Xingyu Wang, Email: xingyu.wang@ndsu.edu, Phone: 701.231.7651

Reporting Period: 12/28/2024 – 3/27/2025

Project Activities for Reporting Period:

In the quarterly report, the teams have changed Principal Investigator (PI), and subcontracts have been renewed. Additionally, we organized and conducted routine biweekly meetings continually. Building on the previous 9^{th} report, the research team focused on tasks 2.2, 3.1, 4.1, 4.2, 5.1, and 6.1 during this quarter (Quarter 10). The following sections provide detailed summaries of the major activities completed during this reporting period.

Task 2.2, Formulate the remaining useful life prediction model: During this reporting period, the research team, consisting of Dr. Zhibin Lin, Dr. Hong Pan, and Mohsin Ali Khan (UTA), developed a framework for remaining useful life (RUL) prediction. A summary of the key activities and findings are provided below:

1) Remailing useful life
The RUL prediction model includes four parts as shown in Figure 1. The process starts with collecting raw data from sensors measuring pressure, temperature, or ultrasonic guided waves (UGW). This data is then pre-processed through denoising, feature extraction, and normalization to create meaningful features. These features are fed into machine learning or deep learning models such as CNNs, LSTMs, Transformers, Autoencoders, or physics-informed hybrids to predict the health of the system. The final output includes RUL predictions and anomaly alerts to support maintenance decisions.

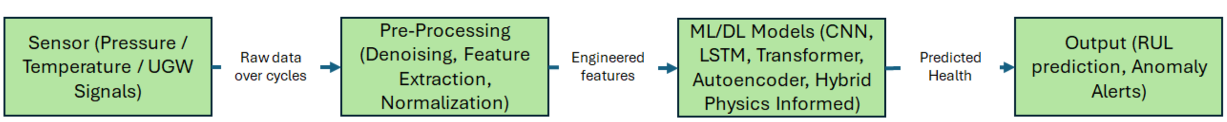

Figure 1: System architecture and data flow for RUL prediction.

2) Feature extraction methodology
Rather than feeding raw signals directly into prognostic models, it is recommended to derive a set of features that contain information relevant to material degradation. Feature engineering is guided by both domain knowledge (fracture mechanics, wave propagation physics) and data-driven

exploration. Table 1 summarized the potential feature extraction methods and the purpose for fracture-based RUL.

Table 1: Preprocessing and Feature Engineering Techniques

Preprocessing Step	Techniques Used	Purpose
Noise filtering	Band-pass, low-pass filters, Spectral subtraction on UGW	Remove high-frequency instrumentation noise and baseline drift; isolate signal bands of interest (e.g. UGW excitation frequency).
Wavelet denoising	Multi-level wavelet decomposition with soft/hard threshold, Adaptive threshold selection (e.g. minimax or SURE criterion)	Suppress random noise in UGW signals while preserving true echo signals. Yields clearer defect indications (higher SNR).
Denoising autoencoder	1D CNN-based autoencoder trained on noise-corrupted signals	Learn complex noise patterns and remove them, which will further improve signal clarity for small defect detection.
Outlier removal	Z-score outlier detection- Physics-based bounds (e.g. pressure cannot exceed design limit)	Eliminate specious data points that could be bias models. Ensure continuity and consistency of time-series.
Elbow point detection	Kneedle algorithm for curve inflection, Segmented regression (piecewise linear fit)	Identify the cycle at which damage progression accelerates. Marks transition to critical degradation phase for early warning.
Time domain feature extraction	Peak, mean, range per cycle (pressure/strain), UGW echo amplitudes, arrival times, Damage index (ratio of defect echo to baseline)	Summarize raw time signals into meaningful indicators. Capture magnitudes of loads and responses that correlate with damage.
Frequency domain feature extraction	FFT spectral centroid, bandwidth- Mode-specific amplitude (e.g. at 50 kHz tone)	Detect changes in frequency content due to damage (e.g. increased attenuation at high frequency). Complements time-domain features.
Time-Frequency feature extraction	Wavelet coefficient energy in particular scales, STFT at damage-sensitive frequencies	Account for non-stationary signal changes and dispersion. Monitors if certain wavelet scales (frequency band) lose energy as crack grows.
Physics-based features	Paris-law model estimate of crack length per cycle, Stress intensity ΔK per cycle- Hydrogen embrittlement factor q	Introduce domain knowledge into data, i.e., approximate how large the crack might be and how aggressive the environment is. Helps models learn physical connection (higher ΔK corresponds to accelerated crack growth).

Preprocessing Step	Techniques Used	Purpose
Feature Scaling	Min-max normalization (0–1), Standardization (mean 0, std 1) per feature	Normalize feature ranges to ensure ML model training stability and that no single feature dominates due to scale.
Dimensionality reduction	PCA on UGW signal set, Autoencoder latent features	Reduce thousands of UGW signal points to a few key features. For instance, the first 3 principal components capturing > 95% of the variance. This compresses data while retaining information for the model.
Data augmentation	Add Gaussian noise to features in training, Synthesize minor variations of cycles via simulation	Increase training sample diversity (especially since failures in real data are rare). This helps prevent overfitting and improves model robustness.

Task 3.1, Preparing the near real-world testbed for hydrogen testing: During this reporting period, the research team, including Mr. J. Anderson from EERC, continues the testing process, with some modification based on safety considerations. The progress is summarized as:

1) The Project received the go-ahead to resume on the EERC side.
2) Reviewed the system design and project scope to determine at this point what needs to be ordered to begin fabrication in the upcoming months. Currently waiting on some insight/clarification to make sure the correct number of materials are specified to ensure costs stay within budget parameters.
3) Upon review of materials, orders will be placed, and fabrication will begin as materials arrive.
4) Design may change slightly due to space limitations and other equipment use in the area. The original pump may not be usable, but a gas booster compatible with hydrogen will be bought instead as we currently use a few in different systems, so we have a vendor in mind.
5) Plan to begin drafting a test plan with NDSU before fabrication is done so after shakedown of the system to ensure its safe operation, the test can begin as soon as possible.
6) Depending on material lead times and EERC operator staff availability, the goal is to be able to have the fabrication of the pipeline done before July, as the month of April and May are already booked near solid, with a little time here and there to do the smaller portions. The biggest hurdle will be the welding and assembly of the pipeline itself, as it will require our certified welder to do it between his other obligations.

Task 4.1 Gaining an understanding of long-term hydrogen impacts, & Task 4.2 understanding of hydrogen adsorption and distribution in existing aged pipe materials through macro-scale simulation: During this reporting period, the Virginia Tech team, led by Dr. K. Wang, focused on hydrogen absorption simulations and finite-element simulation of stress distribution around a pipe elbow. These efforts are summarized as follows:

We have developed a computational model based on Fick's second law for simulating the transport of hydrogen within pipeline components/sections, possibly with pre-stress. Figure 2 shows the computational grid for two example cases, featuring a quarter model (with symmetry boundaries) of a cylindrical pipe section with different wall thicknesses.

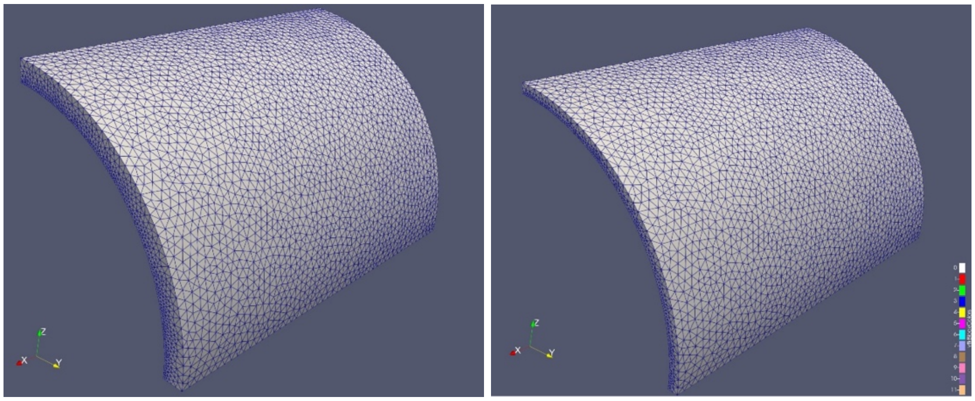

Figure 2: Computational grid for two pipeline sections with different wall thicknesses.

The governing equations being solved is Fick's second law of diffusion, i.e.,

$$\frac{\partial \varphi}{\partial t} - \nabla \cdot (D \nabla \varphi) = 0,$$

where $\varphi(\boldsymbol{x}, t)$ (mol/m3) represents the concentration of hydrogen at any point $\boldsymbol{x}$ within the pipe wall at any time $t > 0$. A Dirichlet boundary condition, $\varphi = \varphi_0$, is specified on the inner wall of the pipe corresponding to the hydrogen concentration in the transported gas, and possibly surface roughness and defects (see previous progress report).

In the equation, D (m^2/s) represents the diffusivity of hydrogen within the pipeline material (e.g., steel). In an idealized case where the pipe material is homogeneous, has no defects, and under no stresses, D can be treated as a constant coefficient. We have used this idealized case to verify our in-house solver, yielding the result shown in Figure 3 below. (A rectangular domain is used in this test case.)

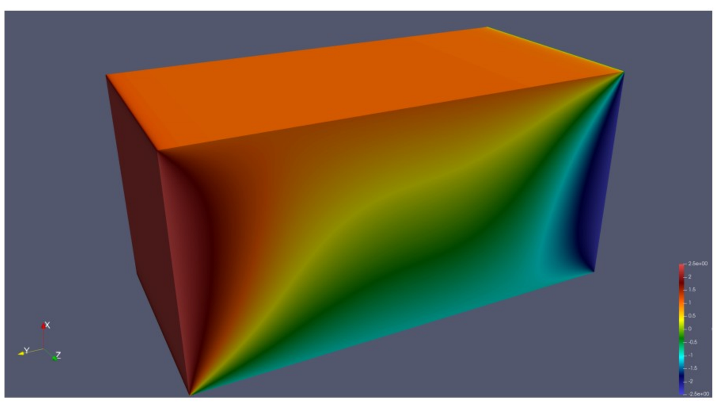

Figure 3: Steady-state solution of Fick's second law with constant diffusivity, D.

In practice, however, D depends on many factors, including but not exclusive to

- **Microscopic material defects**, such as dislocations and grain boundaries. The hydrogen diffusivity tends to be higher within and near these defects.
- **Macroscopic material defects** such as those caused by welding

- **Pre-existing stresses** --- In particular, tensile stresses tend to promote hydrogen diffusion absorption, while compressive stresses tend to hinder hydrogen absorption
- **Material variation** such as near welded joints

Therefore, in our studies, we need to treat D as a variable, accounting for these factors. We have started with considering D as a function of pre-existing stress. For this purpose, we first conducted a finite-element stress analysis for a pipe elbow, as shown in Figure 4 below.

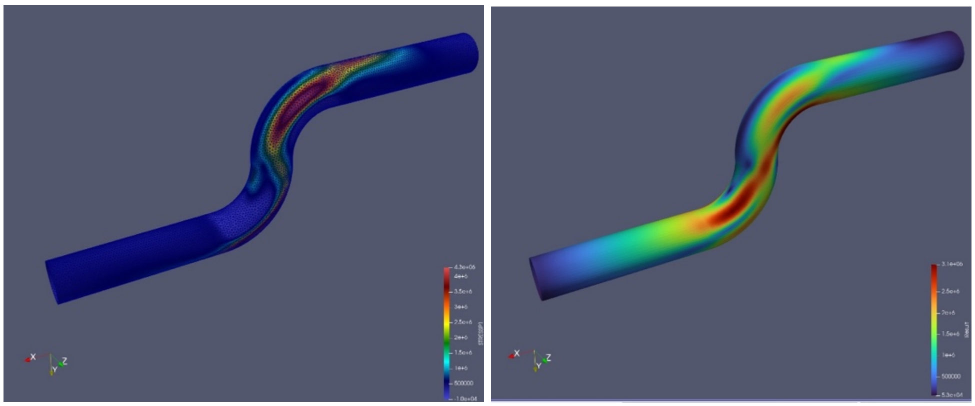

Figure 4: Finite-element simulation of stress distribution around a pipe elbow. Left: Maximum principal tensile stress; Right: Maximum shear stress.

Next, we plan to adopt an empirical model for D as a function of the maximum principal tensile stress obtained from the finite element structural mechanics analysis, an re-run the hydrogen diffusion simulations. Furthermore, we plan to account for the reciprocal effects of hydrogen concentration on the pipe material's mechanical properties --- generally speaking, absorption of hydrogen increases the extent and magnitude of defects, which leads to material degradation (e.g., lower Young's modulus values). This will leads to a weakly two-way coupled analysis in which the hydrogen diffusion simulation and the structural mechanics analysis are conducted iteratively with parameters updated based on the latest results obtained from the other solver.

So far, all the simulations have been conducted using the VT research team's in-house open source solvers: M2C and Aero-S. As a back-up plan, we have also obtained access to COMSOL Multiphysics, which can also conduct some, if not all, of the simulations.

Task 5.1, Develop the fracture based remaining useful life prediction model: During this quarter, the research team, including Dr. Zhibin Lin and Dr. Hong Pan from UTA, and Mohsin Ali Khan from UTA, continued their research on physical informed deep learning models for remaining useful life predictions especially Paris' Law based remaining useful life prediction. Their findings are summarized as follows:

In fracture mechanics, the RUL can be estimated using crack growth variation under cyclic loading conditions. Paris' law gives the crack growth per cycle: $da/dN = C \times (\Delta K)^m$, where a is crack length, N are the number of cycles, ΔK is stress intensity factor range, and C, m are material constants determined empirically. For X52 steel in air, certain C, m under the hydrogen environment, are increased which reflects the accelerated crack growth rate. Researchers have extended Paris' law for hydrogen environments and introduced factors dependent on hydrogen pressure, loading and frequency. An empirical model presented express as:

$$C_{hyd} = C_{air} \times [1 + (4.6 - 4.6 \times e^{-0.05P}) \times \left(3^{2(1+R)} \times f^{-0.08} \times q\right]$$

where P is hydrogen pressure (MPa), R is stress ratio, f loading frequency, and q a hydrogen sensitivity factor. The presented equation modifies the Paris law coefficient C to account for hydrogen effect. The factor q is further related to steel composition and strength i.e., higher yield strength or certain alloy elements can change q.

$$q = 1.072 + 0.00011(YS) - 0.5161(Mn)$$

We aim to incorporate a physics-based RUL estimation. Paris' law modified for hydrogen provided crack growth cycle-by-cycle. An assumed initial flaw size (from fabrication or worst-case undetected crack), we can integrate da/dN until the crack reaches a critical size a_{crit} (where failure by fracture is expected). This could be helpful to provide a physics-estimated life N_{phys}. Because from sensor data, we can continuously update this estimate. For instance, if UGW signals suggest a crack has grown faster than expected, we can adjust the current crack length input. Essentially, the physics model runs in parallel with the ML model. It provides an RUL estimate based on known physics. However, due to model uncertainties (exact 'C' and 'm' for given hydrogen mix, or initial flaw unknown), the physics-alone RUL might be inaccurate. For instance, hydrogen effects might not be fully captured by a single C factor, but multiple regions of crack growth behavior can occur. Therefore, it is recommended that we not rely solely on the Paris law output but use it as a feature and a consistency check.

Task 6.1, Examine the potential of transformer-based model for guidelines/best practices summarization: During this quarter, the research team, including Dr. Zhibin Lin, Dr. Hong Pan, and Mohsin Ali Khan from the University of Texas at Arlington (UTA) investigated the potential of leveraging transformer-based models (e.g., GPT) to summarize and elucidate best practices and guidelines. Their findings are summarized as follows:

Transformer-based models, such as GPT, have demonstrated strong capabilities in processing and understanding large volumes of unstructured text. Their ability to capture long-range dependencies through self-attention mechanisms makes them well-suited for identifying key themes, extracting actionable insights, and generating concise summaries from complex technical documents. When applied to engineering guidelines, these models can automatically distill lengthy reports into structured summaries, highlight critical recommendations, and rephrase domain-specific instructions into more accessible language for diverse stakeholders.

The team explored using pre-trained language models fine-tuned on domain-specific corpora to improve accuracy and relevance. These models can not only identify the best practices across multiple documents but also compare and contrast conflicting recommendations, providing a more coherent interpretation of standard procedures. Moreover, transformer models are capable of organizing extracted information into categories such as safety protocols, operational thresholds, inspection routines, and maintenance schedules—enabling better traceability and decision support. In summary, transformer-based models show strong potential for automating the summarization of best practices and enhancing the clarity and usability of engineering guidelines. This capability can significantly reduce the manual effort required to interpret technical documents, promote standardization, and facilitate knowledge transfer across teams and projects.

Project Financial Activities Incurred during the Reporting Period:

The cost breakdown for each budget category during the reporting period is presented in Table 2. The PI change and subaward process were completed at the end of January; please note that the associated costs were delayed due to administrative processing between institutions.

Table 2 Cost breakdown during the reporting period (Q10)

Category	Amount spent during Q10
Personnel	
Faculty	$0.00
Postdoc	$0.00
Students (RA and UR)	$0.00
Benefits	$0.00
Operating Expenses	
Travel	$0.00
Materials and Supplies	$0.00
Recharge Center Fee	$0.00
Consultant Fee	$0.00
Subcontracts	$25,575.99
Indirect Costs	$0.00

Project Activities with Cost Share Partners:

As previously mentioned, the PI change and subaward process were completed at the end of January. Due to the subcontract suspension, there have been no significant expenses during this period; as a result, the matching funds have also been paused. Moving forward, matching funds will primarily be provided through RA tuition waivers for those continuing work on this project.

Project Activities with External Partners:

During this reporting period, the research team meets regularly bi-weekly, and the sub-universities have researched as planned.

Potential Project Risks:

No potential risks were noticed during this reporting period.

Future Project Work:

During the upcoming quarter, the research team will persist in their efforts on Tasks 2.2, 3.1, 4.1, 5.1, and 6.1, with a specific emphasis on accelerating task 3.1 for near real-world testbed.

Potential Impacts on Pipeline Safety:

The hydrogen absorption simulation results provide a strong foundation for future safety-related hydrogen simulations, offering insights into material behavior under hydrogen exposure. In parallel, deep learning models for remaining Useful life prediction can accurately forecast system degradation. Combined with simulation and experimental results, this approach can enhance predictive maintenance and improve the safety and reliability of hydrogen pipeline infrastructure.