A Machine Learning Emulator for Hydrogen Embrittlement

Minh Triet Chau
s6michau@uni-bonn.de

João Lucas de Sousa Almeida
IBM Research Brazil

Alberto Costa Nogueira
IBM Research Brazil

Abstract

We propose a Physics-informed Machine learning model to emulate hydrogen embrittlement in pipelines by predicting the gas pressure on the pipe's inner wall. Despite its high-fidelity results, the current PDE-based simulators are time- and computationally-demanding. Using simulation data, we train a machine learning model to predict the time and location of a failure much faster with controllable loss in accuracy.

1 Introduction

Climate change requires clean and effective energy storage to replace gasoline, coal, or natural gas (NG). Batteries are a clean carrier but do not have sufficient energy density for sectors such as cement, steel, and long-haul transport [Emma et al., 2021]. For those industries, one option that has received considerable attention is low-carbon hydrogen [McQueen et al., 2020], which can store a large amount of energy and does not release greenhouse pollutants in combustion. However, the inefficiency of green H_2 manufacturing process is one of the biggest obstacles to its dissemination [Joshi et al., 2022]. While finding an environmentally friendly and affordable way to produce H_2 is a long-term task, it should not deny us hydrogen's immediate benefit. One viable strategy is blending H_2 with NG (HCNG) to reduce emissions when burning [Melaina et al., 2013]. When increasing the volume of H_2 from 0% to 15%, [Pandey et al., 2022] observe up to 50% reduction in CO_2 emission. Blends with less than 20% H_2 by volume can be transmitted by repurposing existing NG pipelines, which are 67% cheaper [Peter et al., 2020] than building new ones. However, the biggest problem with repurposed pipelines is during daily consumption, gas pressure may reach excessive values that lead to hydrogen diffusion through the most current pipeline wall materials [EU Agency for the Cooperation of Energy Regulators, 2021]. This phenomenon is known as hydrogen embrittlement (HE). An embrittled pipeline is prone to leaking H_2, which can cause catastrophic events since even small amounts of H_2 can ignite in the presence of air. Therefore, HCNG is not yet mainstream. In Germany, where it is most widely adopted, HCNG only accounts for 10% of demand per capita [Dolci et al., 2019].

Preventing HE requires monitoring, operational pressure management, and pipeline maintenance [Ronevich and San Marchi, 2019]. Since H_2 diffusion strongly correlates with pressure [Hafsi et al., 2018], we propose a Machine Learning model[1] to predict pressure at each point in the pipeline from past data. This prediction let us know where and when the next leakage will happen. The further into the future we can forecast, the more options there are [Raymond, 1988, Castagna and Stoloff, 1995, Liu and Atrens, 2015, Michler et al., 2021] to mitigate the hazardous effects.

[1] Source code is available at `https://github.com/minhtriet/hydrogen_emb`.

Preprint. Under review.

We will organize the paper as follows. Firstly, we present the specifications for our simulation and the challenges we face. Then we discuss the choice of machine learning models to emulate our dynamic system and the details of the experiments. Finally, we discuss about the ML models and their usability in predicting HE.

2 Problem set up

We run a simulation with the specifications in Table 1 to get the raw data. Given these parameters, the Reynolds number of the flow is $Re = \frac{\rho u D}{\mu} = \frac{0.08988 \times 20 \times 7.62e^{-2}}{0.88e^{-5}} \approx 15565.58$. When the Reynolds number is larger than 2900, the flow is turbulent [Schlichting and Gersten, 2016], which makes the problem chaotic and tricky for ML models to extrapolate. For such systems, the most accurate method, as well as time- and computationally-demanding, is solving an approximation of the Navier-Stokes equation. With SimScale simulation (Figure 1a and 1b), a run of 1000 time steps with fixed $\Delta t = 0.002s$ takes 30225 seconds. From a practical perspective, this long simulation time means the incident may occur even before the simulation finishes. From an ML perspective, training a deep neural network with such a small amount of data is prone to overfitting. We will address those concerns in the next session.

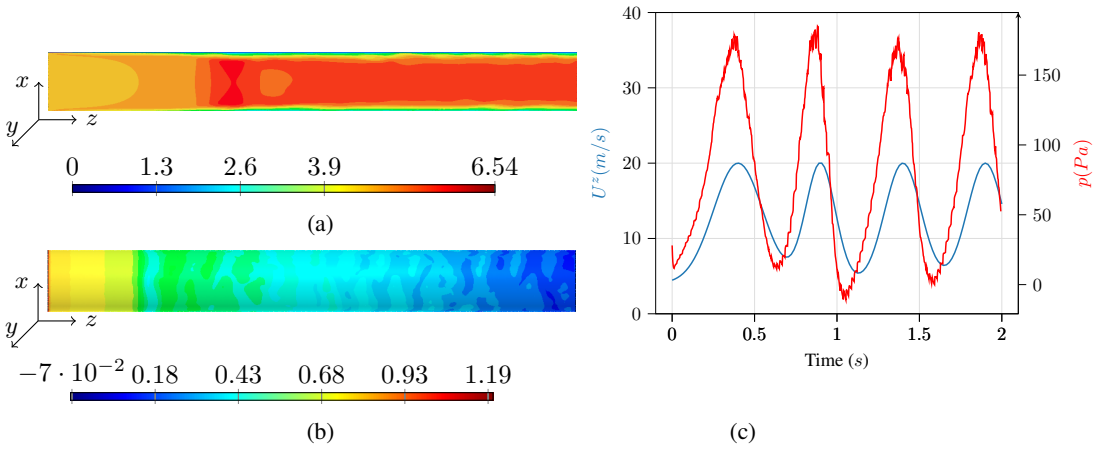


Figure 1: (1a): The velocity field profile $\|\mathbf{U}\|_2$ of gas flow (m/s) in a pipeline. (1b): The pressure field of gas flow to the inner wall. (1c) The inlet velocity U^z and mean inner wall pressure of the whole pipeline through time. The periodic profile for U^z is to imitate the real-life demand.

$$\begin{pmatrix} \mathbf{X}_1 & \mathbf{U}_1 & k_1 & \omega_1 & p_1 & nut_1 \\ \mathbf{X}_2 & \mathbf{U}_2 & k_2 & \omega_2 & p_2 & nut_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{X}_n & \mathbf{U}_n & k_n & \omega_n & p_n & nut_n \end{pmatrix}$$

(a)

$$\text{time } t \downarrow \quad \xrightarrow{\text{pipeline length in axis } z} \quad \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m,1} & p_{m,2} & \cdots & p_{m,n} \end{pmatrix}$$

(b)

$$\begin{pmatrix} k_{1,1} & k_{1,2} & \cdots & k_{1,n} \\ p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ \mathbf{U}_{1,1} & \mathbf{U}_{1,2} & \cdots & \mathbf{U}_{1,n} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{U}_{m,1} & \mathbf{U}_{m,2} & \cdots & \mathbf{U}_{m,n} \end{pmatrix}$$

(c)

Figure 2: The organization of features. Figure 2a: Raw simulation output at position $\mathbf{X} \in \mathbb{R}^3$ - gas velocity $\mathbf{U} \in \mathbb{R}^3$, pipeline internal static pressure p, rate of dissipation of kinematic turbulence k, kinematic turbulence energy ω, kinematic turbulent viscosity nut. Figure 2b: Snapshot matrix for p with the number of pipeline grid points n, and the total number of simulation time steps m. Figure 2c: An example of snapshot stacking using features $(k, p, \mathbf{U})$.

3 Numerical experiments

From simulation output (Figure 2a), we construct snapshots and stack them (Figure 2b and 2c). The first 50% of data is the training set, and the next 10% is the validation set. After that, we retrain the model with the train and validation sets to predict the test set. To reduce the computational cost, we first apply PCA to generate a low-rank representation of the dynamic system. Then, we apply Transformers, Temporal Fusion Transformers (TFT) [Lim et al., 2021] and Operator Inference [Ghattas and Willcox, 2021], a physics-aware method in the reduced order space. The inference step is done in one step. In other words, at the end of the validation set, each algorithm will immediately produce a $\mathbb{R}^{m_{test} \times n}$ matrix, whose semantic is similar to Figure 2b. Regardless of the methods, their inference time is less than a second compared to 30225 seconds in the numerical simulation (cf Session 2).

Parameter	Value
Pipeline	
Diameter (cm)	7.62
Length (cm)	500
Boundary condition (BC)	
Inlet velocity U^z	Detailed in Figure 1c
Relative pressure p_{outlet} (pa)	0
Wall	No slip
Simulation control	
Turbulence model	SST $k-\omega$ [Menter, 1993]
Solver algorithm	PIMPLE
Δt	$0.002s$

Table 1: Simulation setup. The specifications of the pipeline, U^z and p_{outlet} are from [Mohitpour et al., 2003].

Baseline: Transformer and TFT. In this approach, we perform a grid search for the number of encoder and decoder layers, dropout rate, and the number of attention heads. Except for p, the other features are set as covariance series. The training of those models is powered by [Herzen et al., 2022]. Even though TFT performs better than Transformers and OpInf in most cases (Table 2), their prediction (Figure 3a and 3b) show that they may not be the most desirable models.

Operator Inference (OpInf). In its first step, OpInf reduces a full order model to a polynomial representation in latent space as

$$\frac{d}{dt}x_t = c + A(x_t) + H(x_t \otimes x_t) + G(x_t \otimes x_t \otimes x_t) + B(u_t)$$

$\otimes$ operator denotes a column-wise Kronecker product. A, H and G represent linear, quadratic, and cubic state matrices respectively. $B(u_t)$ can be used as the reduced operator of a forcing term in reduced space. c is a constant term. Each of the terms may or may not be used by OpInf, depending on the experimenter's choice. In this work, we use A, H, B operators and c. In the second step, OpInf solves a least square problem to infer these operators.

Two key parameters to optimize for OpInf are the number of the basis for its dimension reduction, and the regularizer for the least square solver since RMSE minimization alone may lead to overfitting.

Number of PCA components. We measure the quality of the reduced order model by computing the preserved energy with respect to the original dynamical system and calculating the RMSE between the original state variables and a PCA reconstruction of those variables. We found out that regardless of the subsets of input $(p, \mathbf{U}, k, nut, \omega)$, 30 basis capture at least 99.82% of the cumulative energy of the system. Based on these criteria, we chose 30 as the number of eigenmodes.

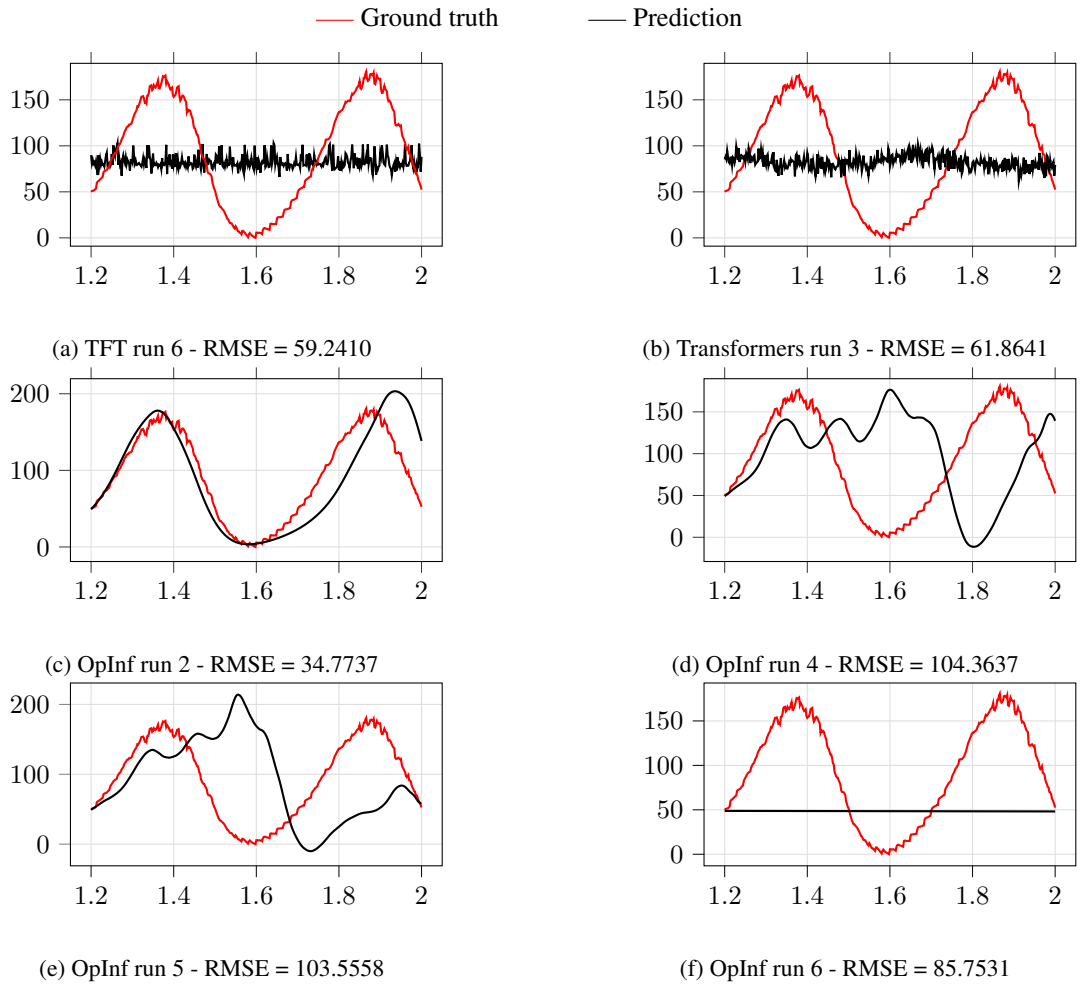
Regularization penalties. Even though the OpInf run 6 (Figure 3f)'s RMSE is smaller than run 4 (Figure 3d) and 5 (Figure 3e), it fails to capture the system dynamics induced by the BC. The reason it does better is that runs 4 and 5 miss the peaks and valleys of p, albeit their wave-like prediction. OpInf run 2 (Figure 3c) gives the best result as it uses the two most important features U^z and p. p is what we are trying to predict, and the flow is strictly dominant in the z-axis (See Figure 1).

4 Conclusion

We propose an ML model for emulating HE through velocity and pressure. Compared to numerical simulations, our ML model takes less than a second to produce a reasonably accurate inference of the pressure in a pipeline. It can help to predict the time and location and time of the next pipeline failure, consequently giving more time for taking action. With it, HCNG is one step closer to mass adoption today and we have better infrastructure for large-scale H_2 delivery in the future.

		RMSE			
Run	Input	Transformer	TFT	OpInf	Regularizer (OpInf only)
1	p	61.2973	**59.1081**	66.5808	910.204
2	p, U^z	60.5701	58.3013	**34.7737**	918.367
3	$p, \mathbf{U}$	61.8641	**59.1939**	78.8812	787.755
4	$p, \mathbf{U}, k$	61.5241	**58.9044**	104.3637	983.673
5	$p, \mathbf{U}, k, \omega$	61.5505	**58.8527**	103.5558	885.714
6	$p, \mathbf{U}, k, \omega, nut$	61.8284	**59.2410**	85.7531	2041734.694

Table 2: Summary for Transformer, TFT, and OpInf approach. See Figure 3 for a visual comparison.



(a) TFT run 6 - RMSE = 59.2410

(b) Transformers run 3 - RMSE = 61.8641

(c) OpInf run 2 - RMSE = 34.7737

(d) OpInf run 4 - RMSE = 104.3637

(e) OpInf run 5 - RMSE = 103.5558

(f) OpInf run 6 - RMSE = 85.7531

Figure 3: Comparison of average pressure to pipeline inner wall through time and its prediction. The x-axis is time (s) and the y-axis is the relative average pressure (pa). Figures 3a and 3b show a typical prediction of Transformers and TFT, regardless of input features and parameter tuning. For OpInf, a higher regularizer used in Figure 3f causes the insensitivity to changes from BC, even though it minimizes test set loss.

References

A Castagna and N S Stoloff. Moisture-induced embrittlement of iron aluminides. final report. 4 1995. doi: 10.2172/93543. URL `https://www.osti.gov/biblio/93543`.

Francesco Dolci, Denis Thomas, Samantha Hilliard, Carlos Fúnez Guerra, Ragnhild Hancke, Hiroshi Ito, Mathilde Jegoux, Gijs Kreeft, Jonathan Leaver, Marcus Newborough, Joris Proost, Martin Robinius, Eveline

Weidner, Christine Mansilla, and Paul Lucchese. Incentives and legal barriers for power-to-hydrogen pathways: An international snapshot. *International Journal of Hydrogen Energy*, 44(23):11394–11401, 2019. ISSN 0360-3199. doi: https://doi.org/10.1016/j.ijhydene.2019.03.045. URL https://www.sciencedirect.com/science/article/pii/S0360319919309693.

Woodward Emma, Han Olivia, and Forbes Charles. Enabling the european hydrogen economy. Technical report, May 2021. URL https://web.archive.org/web/20211008181326/https://www.euractiv.com/wp-content/uploads/sites/2/2021/07/Aurora-MCS-Enabling-the-European-hydrogen-economy-20210322.pdf.

EU Agency for the Cooperation of Energy Regulators. Transporting Pure Hydrogen by Repurposing Existing Gas Infrastructure: Overview of existing studies and reflections on the conditions for repurposing, July 2021. URL https://web.archive.org/web/20220805232342/https://acer.europa.eu/Official_documents/Acts_of_the_Agency/Publication/Transporting%20Pure%20Hydrogen%20by%20Repurposing%20Existing%20Gas%20Infrastructure_Overview%20of%20studies.pdf. publisher: European Union Agency for the Cooperation of Energy Regulators.

Omar Ghattas and Karen Willcox. Learning physics-based models from data: perspectives from inverse problems and model reduction. *Acta Numerica*, 30:445–554, 2021. doi: 10.1017/S0962492921000064.

Zahreddine Hafsi, Manoranjan Mishra, and Sami Elaoud. Hydrogen embrittlement of steel pipelines during transients. *Procedia Structural Integrity*, 13:210–217, 2018. ISSN 2452-3216. doi: https://doi.org/10.1016/j.prostr.2018.12.035. URL https://www.sciencedirect.com/science/article/pii/S2452321618302683. ECF22 - Loading and Environmental effects on Structural Integrity.

Julien Herzen, Francesco Lässig, Samuele Giuliano Piazzetta, Thomas Neuer, Léo Tafti, Guillaume Raille, Tomas Van Pottelbergh, Marek Pasieka, Andrzej Skrodzki, Nicolas Huguenin, Maxime Dumonal, Jan Kościsz, Dennis Bader, Frédérick Gusset, Mounir Benheddi, Camila Williamson, Michal Kosinski, Matej Petrik, and Gaël Grosch. Darts: User-friendly modern machine learning for time series. *Journal of Machine Learning Research*, 23(124):1–6, 2022. URL http://jmlr.org/papers/v23/21-1177.html.

Mohit Joshi, Ilya Chernyakhovskiy, and Mark Chung. Hydrogen 101: Frequently asked questions about hydrogen for decarbonization. July 2022. doi: 10.2172/1879231. URL https://www.osti.gov/biblio/1879231.

Bryan Lim, Nicolas Loeff, Sercan Arik, and Tomas Pfister. Temporal fusion transformers for interpretable multi-horizon time series forecasting. 2021.

Qian Liu and Andrej Atrens. Reversible hydrogen trapping in a 3.5nicrmov medium strength steel. *Corrosion Science*, 96:112–120, 2015. ISSN 0010-938X. doi: https://doi.org/10.1016/j.corsci.2015.04.011. URL https://www.sciencedirect.com/science/article/pii/S0010938X15001717.

Shawna McQueen, Joseph Stanford, Sunita Satyapal, Eric Miller, Ned Stetson, Dimitrios Papageorgopoulos, Neha Rustagi, Vanessa Arjona, Jesse Adams, Katie Randolph, David Peterson, Laura Hill, Mariya Koleva, Tim Reinhardt, Evan Frye, Robert Schrecengost, Angelos Kokkinos, John Litynski, Regis Conrad, Grigorii Soloveichik, David Tew, Scott Litzelman, John Vetrano, Rebecca Onuschak, Alison Hahn, Eric Hsieh, and Ryan Costa. Department of energy hydrogen program plan. 11 2020. doi: 10.2172/1721803. URL https://www.osti.gov/biblio/1721803.

M. W. Melaina, O. Antonia, and M. Penev. Blending hydrogen into natural gas pipeline networks. a review of key issues. 3 2013. doi: 10.2172/1219920. URL https://www.osti.gov/biblio/1219920.

Florian R. Menter. Zonal two equation k-w turbulence models for aerodynamic flows. 1993.

Thorsten Michler, Frank Schweizer, and Ken Wackermann. Review on the Influence of Temperature upon Hydrogen Effects in Structural Alloys. *Metals*, 11(3):423, March 2021. ISSN 2075-4701. doi: 10.3390/met11030423. URL https://www.mdpi.com/2075-4701/11/3/423.

Mo Mohitpour, H Golshan, and A Murray. *Pipeline Design and Construction*. American Society of Mechanical Engineers, Fairfield, CA, July 2003.

Vivek Pandey, Suresh Guluwadi, and Gezahegn Habtamu Tafesse. Performance and emission study of low HCNG fuel blend in SI engine with fixed ignition timing. *Cogent Engineering*, 9(1):2010925, December 2022. ISSN 2331-1916. doi: 10.1080/23311916.2021.2010925. URL https://www.tandfonline.com/doi/full/10.1080/23311916.2021.2010925.

Adam Peter, Heunemann Frank, von dem Bussche Christoph, Engelshove Stefan, and Thiemann Thomas. Hydrogen infrastructure – the pillar of energy transition, 2020.

L Raymond. Hydrogen embrittlement: Prevention and control; astm special technical publication, no. 962. 1 1988. URL https://www.osti.gov/biblio/6715649.

Joseph Allen Ronevich and Christopher W. San Marchi. Hydrogen effects on pipeline steels and blending into natural gas. 10 2019. URL https://www.osti.gov/biblio/1646101.

Hermann Schlichting and Klaus Gersten. *Boundary-layer theory*. Springer, Berlin, Germany, 9 edition, October 2016.