A Machine Learning-Driven Framework for Corrosion Risk Assessment in Long-Distance Pipelines

Li Hengfeng[1,2], Chen Shaosong[3], Li Lin[1,2], **Li Wenwen[1,2*]**, Hu Xiuqian[3], Wang Jinguang[3]

[1.] *Shaanxi Key Laboratory of Shaanxi Province for Gas & Oil Logging Technology, Xi'an, Shaanxi, China 710065*

[2.] *Xi'an Shiyou University, Xi'an, Shaanxi, China 710065*

[3.] *Beijing Ankocorr Technology Co., Ltd., Beijing ,China 102209*

**Presenter's email address: 3157807602@qq.com*

Abstract: The rapid advancement of Industrial Internet of Things (IIoT) and machine learning technologies has exposed limitations in traditional corrosion risk assessment methods, particularly in terms of accuracy and real-time performance. This study presents an innovative machine learning-based framework for corrosion risk assessment in long-distance pipelines. The framework integrates public environmental factors (e.g., meteorological and geological data) with private operation and maintenance data (e.g., intelligent pigging and cathodic protection monitoring data). For data management, customized preprocessing workflows have been designed for various data types, and information is organized in an N-dimensional vector format to ensure data quality and consistency. The modeling component employs an adaptive optimization algorithm based on historical data, incorporating multiple pre-set machine learning models and their hyperparameter spaces. Through automatic adjustment and selection of optimal model configurations, the framework significantly improves the accuracy and generalization capability of risk assessment. Multiple specialized models are coupled using ensemble learning methods, forming an end-to-end risk assessment workflow. This study also explores strategies for model deployment and continuous optimization mechanisms, ensuring the framework's scalability and maintainability in practical production environments. The proposed comprehensive framework aims to enhance the accuracy, efficiency, and adaptability of corrosion risk assessment for long-distance pipelines, thereby providing robust support for operation and maintenance decisions.

Keywords *Corrosion Risk Assessment; Machine Learning; Long-Distance Pipelines*

Reference

[1] Ma Haonan, Zhang Weidong, Wang Yao, Ai Yibo, and Zheng Wenyue, “Advances in corrosion growth modeling for oil and gas pipelines: A review,” *Process Safety and Environmental Protection*, vol. 171, 2023

[2] X. Li, H. Zhao, and R. Zhang, “Data-driven dynamic failure assessment of subsea gas pipeline using process monitoring data,” Process Safety and Environmental Protection, vol. 166, pp. 1–10, Oct. 2022, doi: 10.1016/j.psep.2022.08.006.

[3] M. Feurer, K. Eggensperger, S. Falkner, M. Lindauer, and F. Hutter, “Auto-Sklearn 2.0: Hands-free AutoML via Meta-Learning,” Oct. 04, 2022, arXiv

[4] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, “Optuna: A Next-generation Hyperparameter Optimization Framework,” Jul. 25, 2019, arXiv: arXiv:1907.10902. Accessed: Dec. 22, 2023

[5] Chen T. and Guestrin C., “XGBoost: A Scalable Tree Boosting System,” in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco California USA: ACM, Aug. 2016, pp. 785–794. doi: 10.1145/2939672.2939785.

[6] N. Adabavazeh, M. Nikbakht, A. Amindoust, and S. Ali Hassanzadeh-Tabrizi, “The identification and analysis of pivotal factors influencing the corrosion of natural gas pipelines using fuzzy cognitive map,” Engineering Failure Analysis, vol. 166, p. 108806, Dec. 2024

[7] Huang Yunfei, Qin Guojin, and Yang Ming, “A risk-based approach to inspection planning for pipelines considering the coupling effect of corrosion and dents,” Process Safety and Environmental Protection, vol. 180, 2023