

Maritime Data Quality for Predictive Maintenance: Navigating Efficiency and Reliability

Author: Di Cui¹, Haiyan Xu¹, Kangsheng Zhou², Ke Wang¹, Xiuju Fu¹

¹Institute of High Performance Computing (IHPC), Agency for Science, Technology and Research (A*STAR), 1Fusionopolis Way, Singapore 138632, Republic of Singapore

²Department of Civil and Environment Engineering, National University of Singapore, 21 Lower Kent Ridge Road, Singapore 119077, Republic of Singapore

Maintenance is a crucial aspect of maritime operations, significantly impacting costs and the reliability of equipment and systems by extending their service life. To avoid unexpected downtime and minimize losses, the concept of predictive maintenance (PdM) has emerged. PdM involves predicting unexpected breakdowns or failures of equipment or systems in advance, thereby improving decision-making in maintenance activities and minimizing downtime. PdM integrates historical data, models, and domain knowledge to anticipate pending failures.

Advancements in technology, such as sensorization and the Internet of Ships (IoS) within Shipping 4.0, which emphasizes a shift towards a digitalized, automated, and interconnected maritime ecosystem, have enabled the integration of big data and AI methods into PdM strategies. This transformation has turned PdM from a conceptual idea into an implementable strategy. PdM for ships, particularly for main engines, offers cost savings and enhances sustainability and reliability in ship operations. Ships are equipped with sensors and onboard instrumentation to provide health and condition monitoring, operational assistance, and operational optimization, forming the foundation for PdM capabilities.

High quality data is essential in PdM for ensuring accuracy and confidence in analytical results and decision making. Data quality issues such as missing values, duplicated entries and outliers, without proper refinement, can bias AI methods or even cause breakdowns. Therefore, establishing a PdM data quality framework tailored to the maritime sector is crucial for AI application development. Such a framework would provide standards and guidelines for identifying high-quality data, thereby enhancing the effectiveness of PdM strategies in the maritime industry.

This study applied a tiered framework for assessing data quality. The framework comprises three tiers: the syntactic tier focuses on adherence to predefined standards, the semantic tier concerns the accurate representation of intended meaning considering contextual information, and the pragmatic tier addresses specific application purposes. Each tier measures the data quality from different dimensions, such as consistency, validity, completeness, uniqueness, accuracy, timeliness and accessibility. Emphasizing practical application, the pragmatic tier is explored through the PdM use cases such as anomaly detection and Remaining Useful Life (RUL) prediction. To demonstrate the proposed maritime data quality assessment framework, we utilize a dataset from a tugboat's main engine. The performance of AI methods is improved after data preprocessing according to the proposed data quality assessment framework.

Key words: data quality, predictive maintenance, sensors, Internet of Ships