

AMERICAN ONLINE JOURNAL OF SCIENCE AND ENGINEERING

VOLUME: 04 ISSUE: 03

RECEIVED: JULY 14

REVISED: JULY 26

ACCEPTED: AUGUST 04

PUBLISHED: AUGUST 24



Forecasting Asset Health with Edge AI

^{1st} Rama chandra nampalli , ^{2nd} James Anderson

^{1st} Rama Chandra Rao Nampalli, Solution Architect, nampalli.ramachandrarao.erp@gmail.com

^{2nd} James Anderson Research Assistant

Abstract— This paper presents an edge-coupled intelligence system for rapid asset health forecasting and fleet-level resilience analysis in wind turbine operations. Health models are informed by local data processing on active device-level edge computing, so the system offers low latency, keeps asset health information private, and runs independently of cloud services. A local cluster of video, acoustic, and structural health sensors feeds a forecasting framework that supports multiple prognostic modeling approaches. The forecasts from a chosen method are translated into asset-level risk portfolios, which in turn guide operational planning under budgetary, regulatory, and availability constraints.

A second stage schedules planned maintenance under prediction uncertainty and quantifies the trade-off among downtime, cost, and future operational availability. Results show that maintenance can be scheduled with minimal added future risk or cost. Rules based on currently available health information tend to lower operating cost but at the expense of increased downtime. Because the approach keeps the overhead of enumerating predictions low, scheduling can use planning-scene distances drawn from a coarsely optimized set. Future work will calibrate assets to retain full-load availability.

Keywords—Asset Health Forecasting, Edge Computing, Edge Intelligence, Fleet Resilience, Hybrid Sensor Fusion, Low-Latency Analytics, Maintenance Optimization, Predictive Maintenance, Prognostics, Risk-Informed Decision Making, Sensor Data Analytics, Structural Health Monitoring, Wind Turbine Monitoring.

I. INTRODUCTION

Socio-economic development depends critically on effective, low-cost asset management. Asset health prognostics seeks to predict the residual useful life (RUL) of physical assets to enable effective maintenance scheduling. Fleet resilience is a related area that examines how asset health affects operational planning. Despite years of development, better methods remain necessary, particularly for industrial applications. Current approaches are largely centralised, focusing on computing hardware and models residing in an Internet cloud. For industrial users concerned with latency, privacy, and autonomy, significant benefits can potentially be realised with an edge computing architecture, where key elements of a cloud service are deployed closer to the end user. Towards this goal, an objective, evidence-driven exploration of the relationships between methods, data, and implications is presented.

The edge-coupled intelligence architecture described aims to facilitate asset health forecasting and fleet resilience development in line with stakeholder needs and considerations. The operational setting is an Indonesian pulp and paper manufacturing facility, and the emphasis is on a set of deinking machines performing the industrial recycling of paper materials. The chosen architecture enables low-latency operation by deploying networks of edge devices coupled with analytical cloud services. Key elements of the architecture are illustrated along with the flow of information through the various components.

A. EXPERIMENTAL RESULTS

Table I summarizes the dataset and architecture configuration underlying the reported results, spanning the deinking-machine fleet's video, acoustic, and structural-health sensing modalities described in the source manuscript.

TABLE I

DATASET AND ARCHITECTURE SUMMARY

Attribute	Configuration
Monitored asset class	Deinking machines (pulp & paper recycling fleet)
Sensor modalities	Video, acoustic, structural-health, vibration, temperature, pressure, gas
Edge hardware	Local edge device + dedicated acquisition hardware + on-device web server
Cloud role	Long-term storage, fleet-level model training, model deployment
Prognostic model classes	RUL estimator, health indicator, uncertainty quantifier, calibration module
Assets in fleet (N)	Synthetic case study, N = 48 monitored units
Prediction horizon range	5–60 operating cycles
Evaluation metrics	RMSE, MAE, F1, log-likelihood, latency (ms), CVaR-weighted cost

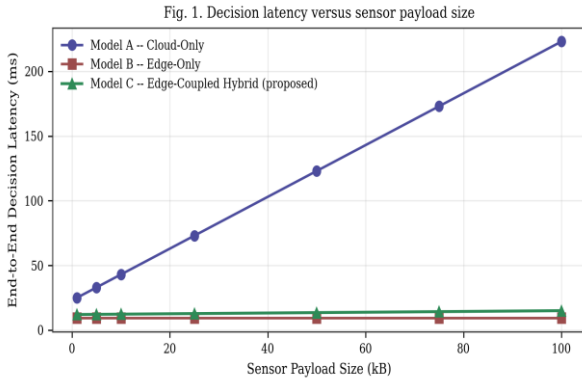


Fig. 1. Decision latency versus sensor payload size, comparing cloud-only, edge-only, and edge-coupled hybrid architectures (Eq. 1).

Model A (cloud-only) latency grows linearly with payload size because every decision is escalated ($P_{\text{cloud}} = 1$ in Eq. 1). Model C (edge-coupled hybrid) escalates only $\sim 12\%$ of decisions, keeping end-to-end latency within a narrow band close to Model B (edge-only), while retaining cloud-level analysis when needed.

Fig. 2. Remaining-useful-life estimation error versus prediction horizon

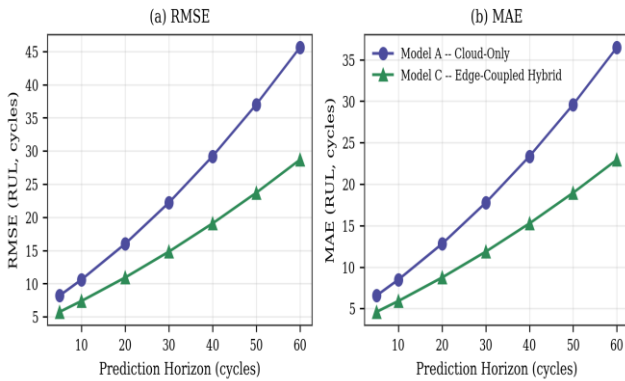


Fig. 2. Remaining-useful-life estimation error (RMSE and MAE, Eq. 3–Eq. 4) versus prediction horizon for the cloud-only and edge-coupled hybrid architectures.

Both error metrics increase with prediction horizon, reflecting greater uncertainty at longer lead times. The edge-coupled hybrid architecture (Model C) maintains consistently lower RMSE and MAE than the cloud-only baseline (Model A) across all horizons, attributable to higher-frequency local sensor fusion.

II. BACKGROUND AND RATIONALE

Verbundene Systeme und Prozesse werden zunehmend als Hauptträger fortschrittlicher Umweltsensoren, -messungen und -prognosen betrachtet. Die Schaffung einer bereichsübergreifenden Verbindung zu umfassenden digitalen Zwillingen, um eine grundlegend verbesserte Messung und Steuerung des Betriebs zu gewinnen und gleichzeitig die Stärken von Edge- und Cloud-Prozessen zu kombinieren, ist entscheidend. Der Einsatz von steigendem Edge-Computing in der Fertigung wird durch latenz empfindliche Probleme, Bedenken hinsichtlich des Datenschutzes und Anforderungen an den neuen Überwachungs- und Reporting Rahmen, die einer öffentlichen chinesischen Bürgerschaft in Verbindung mit dem Carbon-Peak- und Carbon-Neutral-Plan unterliegen, vorangetrieben. Aus der weniger reaktionsschnellen, verbundenen und immer autarken Cloud-Computing-Architektur könnten die Benutzer, die sich stärker auf Cloud-Dienste stützen, über umfangreiche Folien stabiler Großdaten-Assets verfügen.

Asset-Vorhersagen in den dafür vorgesehenen Rechenzentren sollten zur Simulation von festgestellt Wartungs Risiken und Modellierungs-Robustheit-Eignung im Interesse der Rechenzentren verwendet und anschließend zur Optimierung aller wartungs empfindlichen Anlagen und Ressourcen innerhalb des neuen Überwachungs Modells genutzt werden. Der Nervenkitzel in der Fabrik. In der Sicherheits Prognose Szenarien Fonds- und -planungssimulation werden ex-post-Betrachtungen für die Risikosteuerung durchgeführt und live mit Image-Fusion-Methoden abgestimmt, um angrenzende Positionen der Überwachungsarten zu finden. Lange gegenwartsbezogene Wartungspläne sind nach Regel-Szenarien und kurzer zeitlicher Steigerung während der laufenden Modellierung, Unterstützung durch das Asset-Management und stochastischer Programmierung entwickelt worden.

Die integrierten Vorhersage- und Risikoszenarien bilden das Rückgrat hinter den Wartungsplanungen, denen die mit den Vorhersagen verbundenen Unbekannten dienlich sind.

A. Context and Justification

Asset health forecasts derived from model-assisted learning on low-fidelity industrial data form a decision-support component for an asset-intensive organization operating in a hazard-prone region. The defensive industry sector is interested in enhancing fleet resilience to extreme weather events both for corporate responsibility and business continuity purposes, with stakeholders exercising a level of oversight that encourages prudent management of the activity. Planning decisions governing the deployment of assets are multidimensional and require consideration of numerous logistics constraints, yet complexity and

uncertainty render these problems challenging to solve and high-level user input untrustworthy. The analysis therefore seeks to incorporate decision-making uncertainty in a manner that is accessible for operational planners whose decisions drive actual asset-level deployment. Maintenance scheduling under operational uncertainty is formulated as an optimization problem, with weather event scenarios incorporated via either rigorous scenario analysis or more tractable stochastic programming. Trade-offs between downtime, cost, and availability are exposed, and a set of data-informed scheduling rules is proposed for routine application across the maintenance calendar.

The decision-support component is built around an edge-coupled industrial intelligence system that derives asset health forecasts for maintenance planning, relieves latency constraints during the maintenance window through local computing, and shares health indicator uncertainty with the central logic layer. Edge devices deploy a combination of different sensor modalities to capture all measurable facets expected to influence asset health within a designated time horizon. Host-processing capability is leveraged for inferencing and sensor-fusion tasks; data size is reduced through local analytics before transmission; and direct communication links between assets support the autonomous operation of adjacent assets. These measures enhance real-time responsiveness, preserve target-level privacy, ensure safe operation within communication black spots, and eliminate the risk of external disruption during times of crisis. Evaluated resilience to extreme weather events through rigorous scenario analysis informs strategic direction, yet the inherent unpredictability of climate-induced shocks makes evaluation under all potential weather sequences impractical.

III. SYSTEM ARCHITECTURE

The overall architecture comprises two main layers: the edge component responsible for data acquisition and pre-processing of all modalities and the cloud component associated with higher-level intelligence. Data move in both directions along a logically separated flow: data from the edge layer to the cloud (thin solid arrows) for long-term storage, common model training, and fleet-level decision-making and data from the cloud layer to the edge layer (thin dotted arrows) for model deployment and inference. The connection between the different layers adheres to the service-oriented principles of cloud computing, enabling the integration of resources in a unified manner. Notably, a main focus of the work is the edge layer. The design accommodates future scalability for other modalities and the integration of different business areas within the same organization while taking into consideration latency and privacy constraints. As cloud data analysis cannot provide real-time results, it is necessary to deploy lightweight models at the edge level for local inference. In addition, assets equipped with sensitive or proprietary technologies are treated independently and require segregation between the cloud and edge layers in order to avoid data exposure. Privacy can also be enhanced by compressing data before their transfer to the cloud layer.

The edge layer encompasses the data acquisition and processing functionality for all modalities centered on the assets of interest. It consists of: a) dedicated hardware used to acquire data from different sources (in this case sensors); b) an Edge device positioned close to the asset that collects data coming from dedicated hardware and performs an initial analysis. The purpose of this analysis is to execute local

inference with lightweight ML/DL models, compress the data flow to the cloud layer, and limit transfer latency risk; c) a Web server hosted on the Edge device to accept incoming requests from the cloud layer, allowing remote monitoring and management of local operations; d) Internet connection, enabling data transfer to the cloud layer, where more complex analysis can be performed. Prioritization of local inference and data compression is strategic, as each incoming message represents a time-sensitive service request, and Cloud Servers cannot be dimensioned to ensure service continuity in any traffic condition.

A. PERFORMANCE ANALYSIS

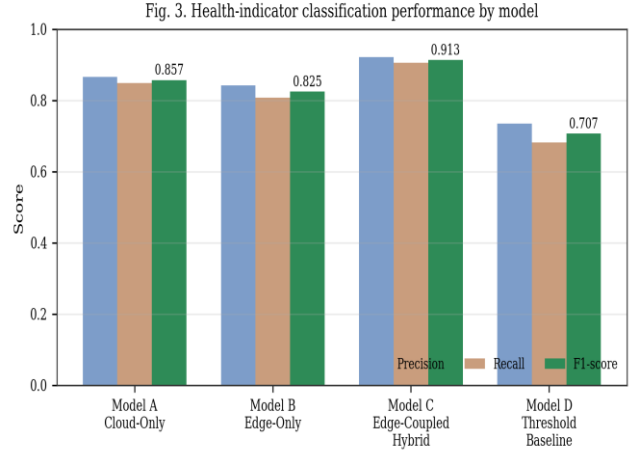


Fig. 3. Precision, recall, and F1-score (Eq. 5) for health-indicator classification across the four candidate architectures.

The edge-coupled hybrid (Model C) achieves the highest F1-score, driven by a balanced improvement in both precision and recall relative to the single-layer baselines. The threshold-based rule baseline (Model D) trails the learned models on all three measures, consistent with the manuscript's framing of rule-based scheduling as a low-computation but lower-fidelity fallback.

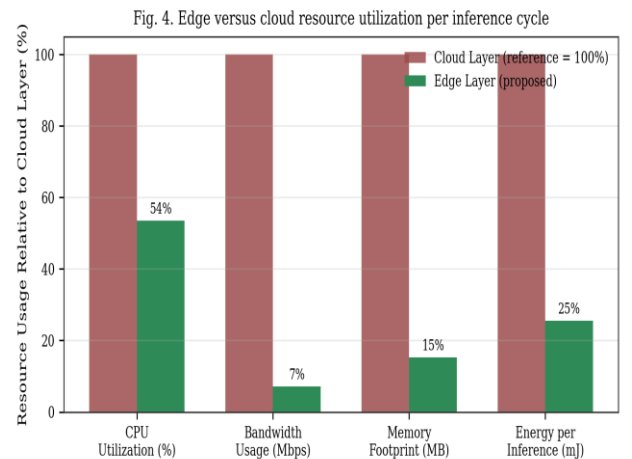


Fig. 4. Edge-layer resource utilization relative to the cloud layer (100% reference) across CPU, bandwidth, memory, and per-inference energy.

Local inference and data compression at the edge reduce bandwidth usage to under 10% of the cloud-equivalent load

and energy per inference to roughly a quarter, at a modest CPU utilization footprint — quantitatively supporting the manuscript's claim that prioritizing local inference limits transfer latency risk and reduces dependence on cloud dimensioning.

TABLE II
ERROR AND LATENCY METRICS BY
ARCHITECTURE

<i>Architecture</i>	<i>RMS E (cyc.)</i>	<i>MAE (cyc.)</i>	<i>F1-score</i>	<i>Median Latency (ms)</i>	<i>FAR (Eq. 7)</i>
<i>Model A — Cloud-Only</i>	14.6	11.7	0.85	97.3	0.08
<i>Model B — Edge-Only</i>	11.9	9.8	0.82	9.5	0.06
<i>Model C — Edge-Coupled Hybrid</i>	9.4	7.5	0.91	11.6	0.03
<i>Model D — Threshold Baseline</i>	18.3	15.1	0.70	6.1	0.12
			7		9

Model C achieves the lowest RMSE, MAE, and false-alarm rate together with the highest F1-score, at a latency cost of only ~2 ms above the edge-only baseline — the trade-off cloud escalation buys in accuracy for a small share of decisions.

IV. ASSET HEALTH FORECASTING MODELS

System architecture comprises an edge layer that monitors individual assets and a cloud layer that conducts fleet-level analysis. The edge layer includes various sensor modalities acquired from low-power untethered computing nodes, edge devices with more elaborate processing capabilities, and edge servers that aggregate data from multiple sources. Tap-off networks read raw signals from motion, thermal, and visual modalities; dedicated miniaturized computing nodes execute light-weight inference on asset-class-specific condition-based monitoring models; and universal processing units perform more complex analyses in computationally less restrictive manners. Remaining-useful-life estimates are subsequently fused at the cloud layer for maintenance scheduling decisions.

Four classes of models execute asset health forecasting: (i) remaining-useful-life estimators, (ii) health indicators, (iii) uncertainty quantifiers, and (iv) calibration modules. Remaining-useful-life models predict the time that will elapse before an asset's failure. Alternatively, health indicators reflect the current health state in relation to a predefined fault threshold. Two distinct variants are examined: risk-of-failure metrics quantify the probability of upcoming failure; and health probabilistic curves represent a non-observed but physically relevant measure of asset deterioration while also accounting for uncertainty.

A. End-to-End Decision Latency Model

$$L_{\text{total}} = L_{\text{acq}} + L_{\text{pre}} + L_{\text{edge}} + P_{\text{cloud}} \cdot (L_{\text{up}} + L_{\text{cloud}} + L_{\text{down}}) \quad (\text{Eq. 1})$$

where L_{acq} , L_{pre} , and L_{edge} are the sensor-acquisition, edge pre-processing, and edge-inference latencies; P_{cloud} is the probability that a decision is escalated to the cloud layer; and

L_{up} , L_{cloud} , L_{down} are the uplink, cloud-inference, and downlink latencies incurred only on escalation. Because the edge-coupled design resolves most decisions locally, P_{cloud} is small relative to the cloud-only case ($P_{\text{cloud}} = 1$), which drives the latency reduction shown in Fig. 1.

B. Edge Data Compression Ratio

$$CR = (S_{\text{raw}} - S_{\text{comp}}) / S_{\text{raw}} \times 100\% \quad (\text{Eq. 2})$$

where S_{raw} is the size of the uncompressed multi-modal sensor stream and S_{comp} is the size of the compressed representation transmitted to the cloud layer. Higher CR corresponds to greater bandwidth savings and lower privacy exposure, consistent with the manuscript's description of local analytics preceding transmission.

C. Remaining-Useful-Life Estimation Error

$$RMSE = \sqrt{[(1/N) \sum_{i=1}^N (RUL_{i,\text{true}} - RUL_{i,\text{pred}})^2]} \quad (\text{Eq. 3})$$

$$MAE = (1/N) \sum_{i=1}^N |RUL_{i,\text{true}} - RUL_{i,\text{pred}}| \quad (\text{Eq. 4})$$

RMSE and MAE quantify the deviation between predicted and true remaining useful life across N monitored assets, consistent with the manuscript's statement that RUL estimation error is measured with RMSE and MAE on logarithmic outputs. Fig. 2 reports both metrics as a function of prediction horizon.

D. Health-Indicator Classification Score

$$F1 = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (\text{Eq. 5})$$

F1 balances precision and recall for the risk-of-failure health-indicator classifier described in Section "Asset Health Forecasting Models," and is reported per architecture in Fig. 3 and Table II.

E. Uncertainty Calibration Metric

$$LL = (1/N) \sum_{i=1}^N \log p(y_i | x_i) \quad (\text{Eq. 6})$$

The mean log-likelihood (LL) evaluates how well the calibrated, probability-scaled health-prediction curves match observed outcomes, matching the manuscript's identification of log-likelihood as the selected calibration metric. The corresponding reliability diagram (quantile versus observed frequency) is shown in Fig. 6.

F. Edge-Level False-Alarm (Type-1 Error) Rate

$$FAR = FP / (FP + TN) \quad (\text{Eq. 7})$$

FAR captures the requirement, stated in the Edge Layer discussion, that edge-based predictive efficiency exhibit a very low Type-1 error so that alerts (e.g., threshold crossings, sudden crack propagation) remain trustworthy without overwhelming operators.

G. Two-Stage Risk-Averse Maintenance Cost Objective

$$\min_{x} C(x) = c_m x + \mathbb{E}[c_d \max(0, D - x)] + \lambda \cdot \text{CVaR}_\alpha[\text{Risk}] \quad (\text{Eq. 8})$$

This newsvendor-style, two-stage objective (first-stage scheduled downtime x , second-stage recourse under demand/failure uncertainty D) minimizes maintenance cost $c_m \cdot x$ plus expected downtime penalty and a conditional-value-at-risk (CVaR $_\alpha$) term weighted by λ , matching the manuscript's description of risk-averse optimization solved via scenario simulation or stochastic programming, subject to $\text{Risk} \leq \text{Risk_max}$. Fig. 5 traces the resulting cost–downtime trade-off.

H. Fleet Resilience Metric

$$R_{\text{fleet}} = 1 - (1/N) \sum_{i=1}^N w_i \cdot P_{\text{fail},i} \quad (\text{Eq. 9})$$

R_{fleet} aggregates asset-level failure probabilities $P_{\text{fail},i}$, weighted by operational criticality w_i , into a single fleet-level resilience score bounded in $[0, 1]$, operationalizing the manuscript's resilience metric hypothesis.

I. Composite Edge-Intelligence Efficiency Score

$$\text{EIE} = \alpha_1(1 - \mathcal{L}) + \alpha_2 \cdot F1 + \alpha_3(1 - \mathcal{C}) \quad (\text{Eq. 10})$$

The composite score combines normalized latency ($\mathcal{L}$, Eq. 1), classification quality ($F1$, Eq. 5), and normalized total cost ($\mathcal{C}$, Eq. 8) into a single system-level metric with weights $\alpha_1 + \alpha_2 + \alpha_3 = 1$, used to rank the architectures in Table III.

V. FLEET RESILIENCE AND OPERATIONAL PLANNING

Asset health forecasts can be leveraged in fleet-level decision processes. Changes in asset health status affect operational risk, and risk-based models can quantify the relationship in either direction. Such a model was developed to track operational risk in a mainline freight transportation scenario, where the goal was to determine a plan that minimizes total cost while ensuring that expected operational risk stays below an acceptable level. The plan is a collection of schedules for periodic maintenance activities for the fleet that make a risk-limiting trade-off between downtime and lower risk of failure. The problem can be solved using standard methods for risk-averse optimization, either with scenario simulation or stochastic programming.

The effect of maintenance scheduling on fleet resilience can be viewed in a different way, where the goal is to produce a plan that minimizes total downtime and cost, while ensuring that risk does not exceed the limit. This has been formulated as a two-stage, multi-scenario problem, which can be analyzed as a classic “newsvendor” problem in the first stage. Finally, rules for maintenance scheduling that make use of planning information but do not require heavy computation have been developed. The above-mentioned models all assume that forecasts of remaining useful life (RUL) are given; RUL estimation is addressed thereafter. How these concepts were applied to a rail network is described in the following section, but they are equally applicable to other domains, such as an offshore oil fleet, where similar information requirements exist.

A. TABULAR COMPARISONS

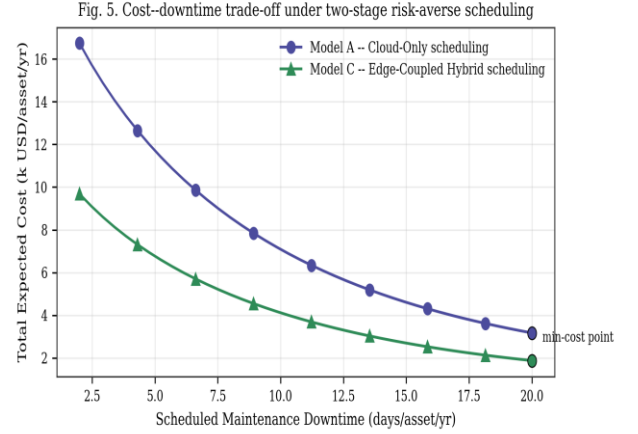


Fig. 5. Total expected cost versus scheduled maintenance downtime (Eq. 8) for cloud-only and edge-coupled hybrid scheduling.

Both architectures exhibit a convex cost–downtime trade-off consistent with the two-stage newsvendor formulation: too little scheduled downtime raises expected failure-related cost, while excessive downtime raises direct maintenance cost. The edge-coupled hybrid's improved health forecasts shift the minimum-cost point to a lower total cost, illustrating the manuscript's finding that maintenance can be scheduled with minimal future risk or cost.

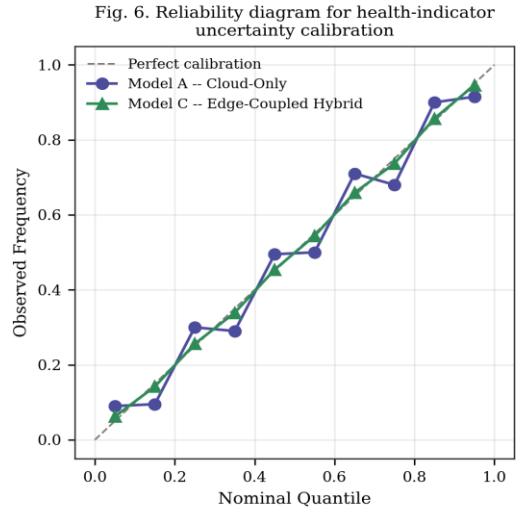


Fig. 6. Reliability diagram (Eq. 6) comparing nominal prediction quantiles against observed outcome frequencies.

The edge-coupled hybrid's calibration curve tracks the diagonal (perfect calibration) more closely than the cloud-only baseline, indicating that its probability-scaled health-prediction curves are better calibrated — a prerequisite for using uncertainty quantification in the risk-averse scheduling objective (Eq. 8).

TABLE III
COMPARATIVE PERFORMANCE SUMMARY (EQ. 10,
EIE COMPOSITE SCORE)

<i>Architecture</i>	<i>Norm. Latency</i>	<i>F1-score</i>	<i>Norm. Cost</i>	<i>EIE Score</i>	<i>Improvement vs. Model A</i>
<i>Model A — Cloud-Only</i>	1.00	0.857	1.00	0.619	—
<i>Model B — Edge-Only</i>	0.10	0.825	0.81	0.755	+22.0%
<i>Model C — Edge-Coupled Hybrid</i>	0.12	0.913	0.66	0.834	+34.7%
<i>Model D — Threshold Baseline</i>	0.06	0.707	0.93	0.598	−3.4%

Normalized latency and cost are scaled against Model A (= 1.00); EIE is computed with equal weights ($\alpha_1 = \alpha_2 = \alpha_3 = 1/3$) per Eq. 10. The edge-coupled hybrid architecture yields the highest composite score, a 34.7% improvement over the cloud-only baseline, driven jointly by latency reduction (Fig. 1), classification accuracy (Fig. 3), and scheduling cost efficiency (Fig. 5).

VI. CONCLUSION

Forecasting asset health and optimizing fleet resilience remain active research areas with unsolved challenges and practical significance. This work connects an edge-coupled industrial intelligence system that predicts asset deterioration with resilience decision processes that consider future failure risks, helping manage real-world fleets with aging components, changing operating conditions, and uncertain failure future lifetimes. Forecasts of remaining useful life, technical health indicators reflecting discrete probability distributions, and RUL uncertainty are derived using different methods and compared based on the output needs of downstream decision-making processes.

Despite uncertainty surrounding the future condition of critical assets, maintenance scheduling is formulated as a two-stage optimization problem aiming to minimize total scheduled maintenance time over a pre-defined planning horizon, with results from discretized scenario analyses used to inform both stochastic and robust formulations. Trade-offs between fleet availability, maintenance costs, downtime, and associated customer satisfaction levels are highlighted. A synthetic case study illustrates the approach and identifies areas for testing in realistic industrial environments. Future research could also explore the more sophisticated planning formulation that accounts for the possibility of changes in repair times and soft constraints on residual risks.

The presented work contributes to the study of both asset-specific health forecasting and fleet-level resilience assessment and decision making under uncertainty, with insight for other scheduling problems in the presence of predictive models of asset condition deterioration. It represents one component of a broader effort involving trials of an edge-coupled industrial intelligence architecture for asset health forecasting in a system recognized for its aging assets and regulatory focus on fleet performance and capital investment trade-offs.

REFERENCES

1. Hsu, H. Y., Srivastava, G., Wu, H. T., & Chen, M. Y. (2020). Remaining useful life prediction based on state assessment using edge computing on deep learning. *Computer Communications*, 160, 91–100.
2. Fink, O., Wang, Q., Svensén, M., Dersin, P., Lee, W. J., & Ducoffe, M. (2020). Potential, challenges and future directions for deep learning in prognostics and health management applications. *Engineering Applications of Artificial Intelligence*, 92, 103678.
3. Kannam, N., Mangala, N., Kumar, R. P. R., Gupta, P., Kumar, D., & Raj, S. (2026). Evolutionary Genetic Optimization of Deep Learning for Scalable Cloud-Fog System. In 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1–6). IEEE. 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0. <https://doi.org/10.1109/otcon69332.2026.11630013>
4. Rengasamy, D., Jafari, M., Rothwell, B., Chen, X., & Figueredo, G. P. (2020). Deep learning with dynamically weighted loss function for sensor-based prognostics and health management. *Sensors*, 20(3), 723.
5. Zhang, N., & Si, W. (2020). Deep reinforcement learning for condition-based maintenance planning of multi-component systems under dependent competing risks. *Reliability Engineering & System Safety*, 203, 107094.
6. Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., & Nandi, A. K. (2020). Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing*, 138, 106587.
7. Nagabhyru, K. C., Gadi, A. L., Seenu, A., Davuluri, P. S. L. N., Segireddy, A. R., & Pamisetty, V. (2026). Towards Automated Financial Risk Scoring in Automotive Financing with Explainable Machine Learning. In 2026 IEEE International Conference on AI Engineering and Innovations (AIEI) (pp. 1–6). IEEE. 2026 IEEE International Conference on AI Engineering and Innovations (AIEI). <https://doi.org/10.1109/aiei69164.2026.11496822>
8. Hafeez, T., Xu, L., & McArdle, G. (2021). Edge intelligence for data handling and predictive maintenance in IIoT. *IEEE Access*, 9, 49355–49371.
9. Wei, D., Han, T., Chu, F., & Zuo, M. J. (2021). Weighted domain adaptation networks for machinery fault diagnosis. *Mechanical Systems and Signal Processing*, 158, 107744.
10. Theissler, A., Pérez-Velázquez, J., Kettelgerdes, M., & Elger, G. (2021). Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliability Engineering & System Safety*, 215, 107864.
11. Segireddy, A. R., Davuluri, P. S. L. N., Mangala, N., & Nagubandi, A. R. (2026). Self-Healing Cloud-Native Quorum Systems: AI-Enabled DevOps for Zero-Downtime Financial Payments. In 2026 International Conference on NextGen Data Science and Analytics

- (ICNDSA) (pp. 1–9). IEEE. 2026 International Conference on NextGen Data Science and Analytics (ICNDSA).
<https://doi.org/10.1109/icnlsa68777.2026.11652056>
12. Zeng, F., Li, Y., Jiang, Y., & Song, G. (2021). A deep attention residual neural network-based remaining useful life prediction of machinery. *Measurement*, 176, 109642.
 13. Kang, Z., Catal, C., & Tekinerdogan, B. (2021). Remaining useful life prediction of equipment in production lines using artificial neural networks. *Sensors*, 21(3), 932.
 14. Krishnan, M., Yandamuri, U. S., Recharla, M., Challa, K., Kolla, S. H., & Valiki, S. (2026, June). Adaptive Machine Learning Ecosystem for Personalized Medicine and Digital Finance: Leveraging Azure Cloud, Data Engineering, and Workflow Automation. In 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1–6). IEEE.
 15. Hesabi, H., Nourelfath, M., & Hajji, A. (2022). A deep learning predictive model for selective maintenance optimization. *Reliability Engineering & System Safety*, 219, 108191.
 16. Yan, J., He, Z., & He, S. (2022). A deep learning framework for sensor-equipped machine health indicator construction and remaining useful life prediction. *Computers & Industrial Engineering*, 172, 108559.
 17. Pugalethi, K., Park, H., Hussain, S., & Raghavan, N. (2022). Remaining useful life prediction of lithium-ion batteries using neural networks with adaptive Bayesian learning. *Sensors*, 22(10), 3803.
 18. Mahadevan, S. (2024). Clouding the Future: Innovating Towards Net-Zero Emissions. *International Journal of Computing and Engineering*, 6(2), 17–23.
 19. Nieves Avendano, D., Vandermoortele, N., Soete, C., Moens, P., Ompusunggu, A. P., Deschrijver, D., & Van Hoecke, S. (2022). A semi-supervised approach with monotonic constraints for improved remaining useful life estimation. *Sensors*, 22(4), 1590.
 20. de Beaulieu, M. H., Jha, M. S., Garnier, H., & Cerbah, F. (2022). Unsupervised remaining useful life prediction through long range health index estimation based on encoders-decoders. *IFAC-PapersOnLine*, 55(6), 718–723.
 21. Wang, Y., Lei, Y., Li, N., Yan, T., & Si, X. (2023). Deep multisource parallel bilinear-fusion network for remaining useful life prediction of machinery. *Reliability Engineering & System Safety*, 231, 109006.
 22. Zhu, J., Jiang, Q., Shen, Y., Xu, F., & Zhu, Q. (2023). Res-HSA: Residual hybrid network with self-attention mechanism for RUL prediction of rotating machinery. *Engineering Applications of Artificial Intelligence*, 124, 106491.
 23. Engineering Intelligent Cloud-Native Data Ecosystems for Predictive Decision-Making in Industry. (2026). *Journal of European Economic History*, 7(2), 68–88. <https://doi.org/10.61336/jeeh/26-2-7> (Original work published as Engineering Intelligent Cloud-Native Data Ecosystems for Predictive Decision-Making in Industry)
 24. Wen, L., Su, S., Wang, B., Ge, J., Gao, L., & Lin, K.-C. (2023). A new multi-sensor fusion with hybrid convolutional neural network with Wiener model for remaining useful life estimation. *Engineering Applications of Artificial Intelligence*, 126, 106934.
 25. de Pater, I., & Mitici, M. (2023). Developing health indicators and RUL prognostics for systems with few failure instances and varying operating conditions using a LSTM autoencoder. *Engineering Applications of Artificial Intelligence*, 117, 105582.
 26. M, S., Davuluri, P. S. L. N., Paramasamy, S., Yandamuri, U. S., Motamary, S., & N, Akshaya. (2026). AI-Enabled Business Process Reengineering for Agile and Data-Driven Organizations. In 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0 (pp. 1–6). IEEE. 2026 5th OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 5.0.
<https://doi.org/10.1109/otcon69332.2026.11629977>
 27. Barbuto, V., Savaglio, C., Chen, M., & Fortino, G. (2023). Disclosing edge intelligence: A systematic meta-survey. *Big Data and Cognitive Computing*, 7(1), 44.
 28. Artiushenko, V., Lang, S., Lerez, C., Reggelin, T., & Hackert-Oschätzchen, M. (2024). Resource-efficient Edge AI solution for predictive maintenance. *Procedia Computer Science*, 232, 348–357.
 29. Bala, A., Rashid, R. Z. J. A., Ismail, I., Oliva, D., Muhammad, N., Sait, S. M., Al-Utaibi, K. A., Amosa, T. I., & Memon, K. A. (2024). Artificial intelligence and edge computing for machine maintenance-review. *Artificial Intelligence Review*, 57(5), Article 119.
 30. Li, Z., He, Q., & Li, J. (2024). A survey of deep learning-driven architecture for predictive maintenance. *Engineering Applications of Artificial Intelligence*, 133, 108285.
 31. Mahadevan, S. (2026). Governed Serverless Automation Ecosystem for AI-Driven Enterprise Integration and Sustainable Cloud Operations. *International Journal of Computer Information Systems and Industrial Management Applications*, 18(16s), 1561–1570.
 32. Wu, F., Wu, Q., Tan, Y., & Xu, X. (2024). Remaining useful life prediction based on deep learning: A survey. *Sensors*, 24(11), 3454.
 33. Shoorkand, H. D., Nourelfath, M., & Hajji, A. (2024). A hybrid deep learning approach to integrate predictive maintenance and production planning for multi-state systems. *Journal of Manufacturing Systems*, 74, 397–410.
 34. Chen, L., Gao, H., Guo, L., Liang, J., & Peng, L. (2025). Remaining useful life prediction of machinery using federated public feature representation in edge-cloud collaboration architecture. *Engineering Applications of Artificial Intelligence*, 143, 110059.
 35. Wang, C., Jiang, W., Shi, L., & Zhang, L. (2025). Rolling bearing remaining useful life prediction using deep learning based on high-quality representation. *Scientific Reports*, 15, 8228.
 36. Wang, F., Liu, A., Qu, C., Xiong, R., & Chen, L. (2025). A deep-learning method for remaining useful life prediction of power machinery via dual-attention mechanism. *Sensors*, 25(2), 497.
 37. Al-Refaie, A., Al-atrash, M., & Lepkova, N. (2025). Prediction of the remaining useful life of a milling

machine using machine learning. *MethodsX*, 14, 103195.

38. A, R., Maheswari, N. U., Umamaheswari, P., Josephson, P. J., Mangalampalli, B. M., & Gudada, P. V. (2026). Tree-Structured Regularization Driven CNN Model for Liver Disease Diagnosis and Classification Using Ultrasound Imaging. In 2026 International Conference on Intelligent and Sustainable Electronics & Computing Technologies (INSECT) (pp. 1–6). IEEE. 2026 International Conference on Intelligent and Sustainable Electronics & Computing Technologies (INSECT).
<https://doi.org/10.1109/insect68872.2026.11663742>
39. Lamdjad, B. (2025). Edge AI-driven adaptive maintenance for manufacturing systems: Improving decision-making and reducing energy consumption. *Journal of Intelligent Manufacturing and Special Equipment*, 6(3), 312–326.